The Technoarete Transactions on Recent Advances in Cyber security and Digital Forensics Journal

Volume No. 12

Issue No. 3

September - December 2023

ENRICHED PUBLICATIONS PVT. LTD

S-9, IInd FLOOR, MLU POCKET,
MANISH ABHINAV PLAZA-II, ABOVE FEDERAL BANK,
PLOT NO-5, SECTOR-5, DWARKA, NEW DELHI, INDIA-110075,
PHONE: - + (91)-(11)-47026006

The Technoarete Transactions on Recent Advances in Cyber security and Digital Forensics Journal

The Technoarete Transactions on Recent Advances in Cyber security and Digital Forensics Journal

(Volume No. 12, Issue No. 3, September - December 2023)

Contents

Security Issues in Web Data Mining, National Security: A Survey

Md Nadeem Ahmed
Research Scholar,
IFTM University, India.

ABSTRACT

Web mining refers to the whole of data mining and related techniques that are used to automatically discover and extract information from web documents and services. When used in a business context and applied to some type of personal data, it helps companies to build detailed customer profiles, and gain marketing intelligence. Web mining does, however, pose a threat to some important ethical values like privacy and individuality. Web mining makes it difficult for an individual to autonomously control the unveiling and dissemination of data about his/her private life. To study these threats, we distinguish between 'content and structure mining' and 'usage mining.' Web content and structure mining is a cause for concern when data published on the web in a certain context is mined and combined with other data for use in a totally different context. Web usage mining raises privacy concerns when web users are traced, and their actions are analyzed without their knowledge. Database mining can be defined as the process of mining for implicit, previously unknown, and potentially useful information from very large databases by efficient knowledge discovery techniques. Naturally such a process may open up new inference channels, detect new intrusion patterns, and raises new security problems. New security concern and research problems are addressed and identified. Finally a particularly well developed theory, rough set theory, is discussed and some potential applications to security problems are illustrated.

Key words: ethics, individuality, KDD, privacy, web data mining

INTRODUCTION

The World Wide Web can be seen as the largest database in the world. This huge and ever- growing amount of data is a fertile area for data mining research. Data mining is the process of extracting previously unknown information from (usually large quantities of) data, which can, in the right context, lead to knowledge. When data mining techniques are applied to web data, we speak of web-data mining or web mining. In accordance with [14], we define web mining as the whole of data mining and related techniques that are used to automatically discover and extract information from web documents and services, based on the definition of Etzioni (1996).

The important ethical issue with data mining is that, if someone is not aware that the information/ knowledge is being collected or of how it will be used, he/she has no opportunity to consent or withhold consent for its collection and use. This invisible information gathering is common on the Web.

Knowledge discovered whilst mining the web could pose a threat to people, when, for instance, personal data is misused, or is used for a purpose other than the one for which it is supplied (secondary use). This same knowledge, however, can bring lots of advantages. Knowledge discovered through data mining is important for all sorts of applications involving planning and control. There are some specific benefits of web-data mining like improving the intelligence of search engines. Web- data mining can also be used for marketing intelligence by analyzing a web user's on-line behavior, and turning this information into marketing knowledge. It should be noted that ethical issues can arise from mining web data that do not involve personal data at all, such as technical data on cars, or data on different kinds of animals. This paper, however, is limited to web-data mining that does, in some way, involve personal data. We shall only look at the possible harm that can be done to people, which means that harm done to organizations, animals, or other subjects of any kind fall beyond the scope of this study. Since most web-data mining applications are currently found in the private sector, this will be our main domain of interest. So, web-data mining involving personal data will be viewed from an ethical perspective in a business context. We clearly recognize that web-data mining is a technique with a large number of good qualities and potential. Web-data mining is attractive for companies because of several reasons. For instance, to determine who might be a new customer by analyzing consumer data, government records, and any other useful information. In the most general sense, it can contribute to the increase of profits be it by actually selling more products or services, or by minimizing costs. In order to do this, marketing intelligence is required. This intelligence can focus on marketing strategies and competitive analyses or on the relationship with customers. The different kinds of web data that are somehow related to customers will then be categorized and clustered to build detailed customer profiles. This not only helps companies to retain current customers by being able to provide more personalized services, but it also contributes to the search for potential customers. So, it is beyond dispute that web-data mining can be quite beneficial to businesses. To make sure that this technique will be further developed in a properly throughout way, however, we shall focus on the possible objections to it. Awareness of all the possible dangers is of great importance for a well-guided development and a well- considered application. Dangers of web data mining lie in the different ways in which privacy is threatened.

DATA MINING AS A SECURITY CONCERN

Thomas Hinke presented an overview of their NASA data mining research. This research does not have any security concerns; in fact it has just the opposite objective to provide content-based metadata (data about data) for the anticipated vast data holding of the Earth Observing System Data and Information System (EOSDIS). This content based metadata would then be used to assist scientists in finding data of interest from the EOSDIS data holdings. However, even though this project has no

provides a useful example of data mining in a scientific-data domain. The problem addressed by the UAH data mining research is that the amount of data in scientific data archives is growing. Some estimates project that projects such as EOSDIS will ingest up to one terabyte of data per day Scientists need to be able to find data of interest -- proverbial ``needle in a haystack.'' The problem of finding data is difficult since there is a lack of content-based metadata. The typical metadata available in the currently operation Version EOSDIS system is limited to satellite, sensor and date captured. All of these represent non content-based metadata. The only content-based metadata available on some data sets are browse images that provide a low resolution view of one of the channels of the data. However, this cannot be automatically processed. From our discussions with users of data at the Marshall Space Flight Center's Distributed Active Archive Center (the EOSDIS archive), there is a general lack of agreement as to what metadata is desired. Thus, this precludes the capture of content- based metadata during data ingest.

INFERENCE PROBLEM AND DATA MINING

Data mining (DM) is an attempt to answer the long-standing question "what does all this data mean?". Such investigations are inherently an attempt to automate the "inference problem" in database security. Inference is basically the process of establishing relationships between datasets, the same objective as data mining. That is, given that certain attributes apply to a set of data, we "know" that certain other attributes also apply to that set of data. This is equivalent to stating that one set "implies" the other. Now, in a multi-level secure (MLS) database, we do not want Low-classified data to infer High-classified data. Data mining processes cannot be used to compromise such rules, of course. This is because each DM process must operate at a specified level (i.e. Low) and must have access to the High data in order to "discover" the rule. However, such Low-to-High rules may be "common knowledge" but unknown to the database designer. Data mining could then be used to combine Low information until the tail of the common-knowledge rule is derived. This is the process of inference. Data are put together "in a surprising way" until some common- knowledge rule, relating Low and High data, can be applied.

Fortunately, data mining can be used effectively to enforce security. The most straightforward way is to search for rules relating Low and High data. We need not be concerned with chains of N inferences, merely what conjunction of attributes for a High set may be implied by Low classified attributes for that set. The security officer doing this analysis has some advantages over an attacker, since he/she has access to both the High and Low data. In most systems, there is relatively little High data, so the number of rules relating High data to Low data is much fewer than the total number of possible rules.

DATA MINING AND SECURITY

Data mining is the process of posing a series of appropriate queries to extract information from large quantities of data in the database. Data mining techniques can be applied to handle problems in database security. On the other hand, data mining techniques can also be employed to cause security problems. This position paper reviews both aspects Data mining techniques include those based on rough sets, inductive logic programming, machine learning, and neural networks, among others. Essentially one arrives at some hypothesis, which is the information extracted, from examples and patterns observed. These patterns are observed from posing a series of queries; each query may depend on the response obtained to the previous queries posed.

Data mining techniques have applications in intrusion detection and auditing databases. In the case of auditing, the data to be mined is the large quantity of audit data. One may apply data mining tools to detect abnormal patterns. For example, suppose an employee makes an excessive number of trips to a particular country and this fact is known by posing some queries. The next query to pose is whether the employee has associations with certain people from that country. If the answer is positive, then the employee's behavior is flagged. While the previous example shows how data mining tools can be used to detect abnormal behavior, the next example shows how data mining tools can be applied to cause security problems. Consider a user who has the ability to apply data mining tools. This user can pose various queries and infer sensitive hypothesis. That is, the inference problem occurs via data mining. There are various ways to handle this problem. One approach is as follows. Given a database and a particular data mining tool, apply the tool to see if sensitive information can be deduced from the unclassified information legitimately obtained. If so, then there is an inference problem. There are some issues with this approach. One is that we are applying only one tool. In reality, the user may have several tools available to him. Furthermore, it is impossible to cover all ways that the inference problem could occur.

PRIVACY THREATENED BY WEB-DATA MINING

In this section, we shall point out that web-data mining, which involves the use of personal data of some kind, can lead to the disruption of some important normative values. One of the most obvious ethical objections lies in the possible violation of peoples' (informational) privacy. Protecting the privacy of users of the Internet is an important issue. Our understanding of privacy, however, is conceptually fragile. The term 'privacy' is used to refer to a wide range of social practices and domains [13]. In this article, we will not discuss the philosophical and legal discussions on privacy. Here, we will restrict ourselves with an informal (and common) definition of informational privacy. Informational privacy mainly concerns the control of information about oneself. It refers to the ability of the individual to

protect information about himself. The privacy can be violated when information concerning an individual is obtained, used, or disseminated, especially if this occurs without their knowledge There are some differences between privacy issues related to traditional information retrieval techniques, and the ones resulting from data mining. The technical distinction between data mining and traditional information retrieval techniques does have consequences for the privacy problems evolving from the application of such techniques [11]. While in traditional information retrieval techniques one has to 'talk' to a database by specifically querying for information, data mining makes it possible to 'listen' to a database (cf. Holsheimer 1999). A system of algorithms searches the database for relevant patterns by formulating thousands of hypotheses on its own. In this way, interesting patterns can be discovered in huge amounts of data. Tavani (1999a) argues that it is this very nature of data mining techniques that conflicts with some of the current privacy guidelines as formulated by the OECD.

ARGUMENTS IN DEFENCE OF WEB-DATA MINING

All the benefits obviously show that web-data mining is a highly valuable technique, which is being developed and applied on a large and growing scale. However, the threats to some important values tend to be rather serious, and will create tension in the web data mining field. Unfortunately, many professionals applying web-data mining in a business context do not foresee any moral dangers in web-data mining. To gain some insight into current web- data mining practices and the attitude of web data miners to the ethical issues involved, twenty of these professionals were interviewed. These interviews combined with a literature study teach us that people prefer to focus on the advantages of web-data mining instead of discussing the possible dangers. Moreover, they revealed several different arguments to support the view that web-data mining does not really pose a threat to privacy and related values. The arguments given in their defense of almost unlimited use of data mining can be sorted into six arguments, and are enlightening. We shall discuss these arguments briefly to show that these arguments do not justify unlimited use of data mining.

POSSIBLE SOLUTIONS

There are means to solve some problems with respect to privacy in the ethical context of web-data mining. We can distinguish solutions at an individual and at a collective level. With solutions at an individual level, we refer to actions an individual can take in order to protect himself/herself against possible harms. For example, using privacy enhancing technologies (PETs), being cautious when providing (personal) information on-line, and checking privacy policies on web sites. The solutions at a collective level refer to things that could be done by society (government, businesses, or other organizations) to prevent web-data mining from causing any harm. For example, further development of PETs, publishing privacy policies, web quality seals, monitoring web mining activities, legal

measures, creating awareness amongst web users and web data miners, and debating the use of profiling. A mixture of technical and non-technical solutions at both the individual and the collective level is probably required to even begin solving some of the problems presented here. But, to what extent can the problems really be solved in both web-data mining categories.

REFERENCES

1. *Database Security IX Status and Prospects Edited by D. L. Spooner, S. A. Demurjian and J. E. Dobson ISBN 0 412 72920 2, 1996, pp. 391-399.*
2. *Pawlak, Z. (1990). Rough sets. Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, 1992.*
3. *Lin, T. Y. (1994), "Anamoly Detection -- A Soft Computing Approach", Proceedings in the ACM SIGSAC New Security Paradigm Workshop, Aug 3-5, 1994, 44-53. This paper reappeared in the Proceedings of 1994 National Computer Security Center Conference under the title "Fuzzy Patterns in data".*
4. *Lin, T. Y. (1993), "Rough Patterns in Data-Rough Sets and Intrusion Detection Systems", Journal of Foundation of Computer Science and Decision Support, Vol.18, No. 3-4, 1993. pp. 225- 241. The extended version of "Patterns in Data-Rough Sets and Foundation of Intrusion Detection Systems" presented at the First Invitational Workshop on Rough Sets, Poznan-Kiekrz, September 2-4. 1992.*
5. *M.J.A. Berry and G.S. Linoff. Mining the Web: TransformingCustomer Data. John Wiley & Sons, New York, 2002.*
6. *R. Clarke. 'Profiling' and Its Privacy Implications. Privacy Law & Policy Reporter, 1: 128, 1994.*
7. *R. Clarke. Platformfor Privacy Preferences: A Critique. Privacy Law & Policy Reporter, 5(3): 46–48, 1998.*
8. *B. Custers. Data Mining and Group Profiling on the Internet. In A. Vedder, editor, Ethics and the Internet, pages 87–104.*
9. *O. Etzioni. The World Wide Web: Quagmire or Gold Mine? Communications of the ACM, 39(11): 65–68, 1996.*
10. *D.R. Fordham, D.A. Riordan and M. Riordan. Business Intelligence. Management Accounting, 83(11): 24–29, 2002.*
11. *J.S. Fulda. Data Mining and Privacy. In R. Spinello and H. Tavani, editors, Readings in CyberEthics, pages 413–417. Jones and Bartlett, Sudbury MA, 2001.*
12. *D.G. Johnson. Computer Ethics, 3rd. edition. Prentice-Hall, Upper Saddle River New Jersey, 2001.*
13. *J.F. Johnson. Immunity from the Illegitimate Focused Attention of Others: An Explanation of our Thinking and Talking about Privacy. In A. Vedder, editor, Ethics and the Internet, pages 49–70. Intersentia, Antwerpen Groningen Oxford, 2001.*
14. *R. Kosala, H. Blockeel and F. Neven. An Overview of Web Mining. In J. Meij, editor, Dealing with the Data Flood: Mining Data, Text and Multimedia, pages 480–497. STT, Rotterdam, 2002.*

The Requirement for Security in E-commerce

Shahin Samimi, Fariba Rastegar
Department of Computer Engineering,
Behbahan Branch, Islamic Azad University, Behbahan, Iran

ABSTRACT

Organizations and companies with given the important and value of their information needs are designing a robust Information Security Management System needs are designing a robust Information Security Management System that Along with environmental changes may need to update your system. Because of advances in technology, enabling organizations to easily access the information and data provided. Firstly, we will introduce the concepts of electronic security and safety protocols and usage of these protocols.

Key words: Electronic commerce, information security, encryption, safety protocols.

1. INTRODUCTION

Lack of firsthand information on real cases has made planning and overcoming security threats much more difficult. Now there are some correspondents got specialty in the field of cyber security, and have many solutions to protect electronic business technologies against potential criminals within cyberspace. Many companies have found that to succeed in electronic business, in addition to security approaches designed to protect sources of information technology, some investments and planning are required to create a comprehensive security program.

2. DIFFERENT ASPECTS OF INFORMATION SECURITY

Today, information security is not considered a new subject. But according to technology development and evolution of data transferring system, methods of protection have got basic and fundamental changes. Generally, security within electronic business may be regarded as a simple relation between data and information protection against internal and external misuses during any stage of electronic business (including register, send, receive, etc.).

Considering these issues within electronic business in general and within electronic funds transfer in particular, has a special situation in designing and developing systems. Such systems should be responsive to security issues which will be explained below.

2.1. Accessibility

A safe and secure system should provide access and availability of data in an appropriate time and place coupled with a protection against any unauthorized access to data. But any system can face the following risks:

The risks include network error, power failure, operational mistakes, application mistakes, hardware error, software system error, and viruses.

The approaches to overcome the risks consist in selecting a preferred communication path, preventing power failure, quality test for software and hardware, limiting access, and providing data support system [1].

2.2. Confidentiality

Confidentiality consists in protecting messages against misuse, tracking, and eavesdropping. Confidentiality can encounter some risks such as unauthorized access by intra-organizational people, and hirelings or by tracking during transmission. Cryptography of messages is usually used to overcome these risks.

2.3 Message Integrity

What it meant by integrity is to prevent any manipulation or unwanted deletion of message. In addition, it includes sequence integrity in order to preventing repetition, and loss of message. Message integrity confronts the risk for incidence of accidental mistakes or those resulting from manipulation within data recording phase, and also output degradation. In order to overcome such risks, the approach to confirm message end to end and message sequence may be used.

2.4 Validity and reliability of message

Another aspect of data systems security is validity and reliability of the message. It consists in providing security for sender and receiver identity, and possibility of affirming transmission and receipt of the message. Validity of message faces the risk of impersonation. Confirming message authentication by a combination of what user knows, what user has, and physical features of user may be used to overcome impersonation [2].

2.5 Inspection and handling capability

It consists in registration of data to be inspected, based on pre-determined conditions for confidentiality, and integrity. Risks and strategies to overcome the risks are the same as mentioned for confidentiality and integrity.

3. ECONOMICALITY FOR DATA SYSTEMS TECHNOLOGIES

As the approaches to overcome the risks threatening security are numerous and various, one of the most important and basic factors having a significant role in providing security for systems, is cost consideration. In other words, a balance should be made between potential risks and cost for overcoming those risks. So, we should not always look for those systems that provide maximum security because such systems occasionally cost a lot and using them is not economically reasonable [3].

4. ENCRYPTION

Encryption includes much of business electronics security. One of the most effective methods to protect network security is encryption of all data that is flowing within the network, and replacing basic words (a simple algorithm form) is the basis for cryptographic algorithm also may be used in digital information. Another common usual method and technique is by using an algorithm equipped with a code key that is actually a series of numbers and consists of encryption rules. Encryption is usually divided to two categories by key management [4].

4.1. Symmetric encryption/Series/Private key

In this method, encryption security depends on a determined and divided code and it is used to encrypt and decrypt at the beginning and the end of the message. The characteristic# feature in this method is that both parties of business exchanges must use the same key for encryption and decryption of electronic data interchange. If the message for electronic data interchange encrypted by the same encryption key, it could not be decrypted by any different key.

4.2. Asymmetric encryption/ Public key (PKI)

This method has been invented to remove symmetric key deficits. A pair of key is used in this method. Each of two keys could encrypt information decryption of which is only possible by another key. A pair of key only assigned to one business partner.

5. PROTOCOLS AND SOFTWARE FOR ENCRYPTION

Various methods of encryption have illustrated above. According to the above mentioned methods, there are numerous software to provide security for electronic exchanges especially within electronic business.

5.1. Data encryption Standard (DES)

This encryption has been suggested by US Department of National Standards. It is an encryption plan

by using a symmetric key. DES uses an alpha number sequence as key to encrypt and decrypt a message. DES is used as hardware in most of computer-based data- processing system. It has a 56 bit key, the processor needs a high speed, and a low speed processor could be applicable.

5.2. Encryption technique by a general and private key (RSA)

In order to solve some of the problems of DES, asymmetric technique RSA introduced. In this method, a private key and a corresponding public key are used instead of a private key for encryption and decryption of messages. Although data protection implemented in the best way by this method, but it is not helpful in identity confirmation.

5.3. Pretty Good Privacy (PGP)

This method is a combination of IDEA and RSA. PGP is either considered a standard for encryption and decryption, or it is the name of a software product for electronic mail. PGP may be used for creating digital signatures by encrypting characters added to the end of the message.

5.4. Digital certificate

The way digital certificate functions is that a person or an organization is going to send an encrypted message applies a digital certificate from an organization involved in issuing digital certificate. Then the relevant organization issues an encrypted digital certificate consisting of a public key for applicant, and other information related to him/her. Of course, the reference issuing digital certificate would provide its own public key publicly.

After sending message, receiver will decrypt it by using a public key enclosed to message, and he will consider whether it has been issued by the reference issuing certificate, then he will get the sender's public key and identity information included within the certificate.

5.5. Security Sachets Layer (SSL)

One the most famous methods to provide security for electronic exchanges within the internet is SSL. SSL is a protocol for encryption produced by Netscape, and it is accepted by most of great producers of internet products. SSL is established too sent confidential documents within the internet; it uses a private key to encrypt sent messages.

SSL encrypts all data exchanged between host and customer. According to process of SSL, rate of function will be reduced considerably.

5.6. Sate Electronic Transactions (SET)

SET is a special protocol designed for bank operations, and transactions by credit cards. SET is an open standard for processing credit cards transactions within the internet which has been invented in cooperation of the great companies producing software such as Microsoft, Netscape, and the great companies for credit cards such as visa, and master card. SET observes confidentiality of transactions in a way that seller has access to demanded commodity information, the price, and whether payment is confirmed, but he has not access to customer payment method.

6. CONCLUSION

Unwillingness to providing information about security deficits rises from a common fear that the public awareness of such deficits can cause customers distrust of the company ability in protecting its own assets, therefore, the company will lose its customers and as a result, it will lose its profitability. As customers do not trust to record financial information online, the companies may not get anything by confirming, voluntarily the fact that each become victims for security-related crimes. Because of media excitements regarding Internet and its capabilities, keeping a positive view in public opinion towards electronic business security is a main concern for most of the companies and it is quite necessary to remain in competition.

REFERENCES

[1] Ravi Kalakota, Andrew B. Whinston. "Electronic Commerce: A Manager's Guide", Addison-Wesley, ISBN: 0-201-88067-9

[2] Davies, Simon G. 1997. "Re-Engineering the Right to Privacy: How Privacy Has Been Transformed from a Right to a Commodity". In Technology and Privacy: The New Landscape Edited by P. Agre and M. Rotenberg. 143-165. Cambridge, MA: MIT Press.

[3] Chaum, David. 1985. "Security Without Identification: Transaction Systems To Make Big Brother Obsolete". Communications of the ACM, 28 : 1030-1044

[4] Steve H. Weingart. "Physical security for the ABYSS system". In Proceedings of the IEEE Computer Society Conference on Security and Privacy, pages 52–58, 1987.

A Brief Survey of Various Security Techniques in Manets

Prachi Garg
Assistant professor in Computer Science,
Geeta Institute of Technology and Management, Kanipla, Kurukshetra.

ABSTRACT

Mobile Ad-hoc Networks (MANET) is an emerging area of research. Most current work is centred on routing issues. A mobile ad hoc network (MANET) is a spontaneous network that can be established with no fixed infrastructure. This means that all its nodes behave as routers and take part in its discovery and maintenance of routes to other nodes in the network .In addition, as ad hoc networks are often designed for specific environments and may have to operate with full availability even in difficult conditions, security solutions applied in more traditional networks may not directly be suitable for protecting them. A short survey over papers on ad hoc networks shows that many of the new generation security techniques are not yet able to address the security problems. To become commercially successful the technology must allow network to support many users. A complication is that addressing and routing in ad-hoc networks does not scale up easily as in the internet.

INTRODUCTION

In view of the increasing demand for wireless information and data services, providing faster and reliable mobile access is becoming an important concern. Nowadays, not only mobile phones, but also laptops and PDAs are used by people in their professional and private lives. These devices are used separately for the most part that is their applications do not interact. Sometimes, however, a group of mobile devices form a spontaneous, temporary network as they approach each other. This allows e.g. participants at a meeting to share documents, presentations and other useful information. This kind of spontaneous, temporary network referred to as mobile ad hoc networks (MANETs) sometimes just called ad hoc networks or multi-hop wireless networks, and are expected to play an important role in our daily lives in near future.

A traditional mobile network consists of a fixed network of servers and clients, with a collection of mobile clients that move throughout the geographic area of the network. Within the mobile network, servers have unlimited power and communicate with mobile hosts over a wireless connection. Mobile clients may only communicate among themselves through a server. Among the issues in this type of network are client power consumption, connectivity of the network, and reachability of mobile clients from a server.

In contrast, a MANET is a collection of mobile servers and clients. All nodes are wireless, mobile and

battery powered [9]. The topology can change frequently. The nodes organize themselves automatically, and can be a standalone network or attached to a larger network, including the Internet [2]. All nodes can freely communicate with every other node.

Ad hoc networks may be very different from each other, depending on the area of application. For instance in a computer science classroom an ad hoc network could be formed between students' PDAs and the workstation of the teacher. In another scenario a group of soldiers is operating in a hostile environment, trying to keep their presence and mission totally unknown from the viewpoint of the enemy. The soldiers in the group work carry wearable communication devices that are able to eavesdrop the communication between enemy units, shut down hostile devices, divert the hostile traffic arbitrarily or impersonate themselves as the hostile parties. As can obviously be seen, these two scenarios of adhoc networking are very different from each other in many ways: In the first scenario the mobile devices need to work only in a safe and friendly environment where thenetworking conditions is predictable. Thus no special security requirements are needed. On the other hand, in the second and rather extreme scenario the devices operate in an extremely hostile and demanding environment, in which the protection of the communication and the mere availability and operation of the network are both very vulnerable without strong protection.

As ad hoc networking somewhat varies from the more traditional approaches, the security aspects that are valid in the networks of the past are not fully applicable in ad hoc networks. While the basic security requirements such as confidentiality and authenticity remain, the ad hoc networking approach somewhat restricts the set of feasible security mechanisms to be used, as the level of security and on the other hand performance are always somewhat related to each other. The performance of nodes in ad hoc networks is critical, since the amount of available power for excessive calculation and radio transmission are constrained, as discussed e.g. in [3]. In addition, the available bandwidth and radio frequencies may be heavily restricted and may vary rapidly. Finally, as the amount of available memory and CPU power is typically small, the implementation of strong protection for ad hoc networks is non-trivial.

The main objective of this paper is to describe how the different security techniques help in various situations in ad-hoc networks. In this firstly the different concepts of manets are discussed related to security then the different techniques for preventing networks are discussed.

VARIOUS ISSUES RELATED TO SECURITY

Physical Security

In ad hoc networks especially mobile nodes are typically significantly more susceptible to physical attacks than wired nodes in traditional networks. However, the significance of the physical security in the overall protection of the network is highly dependent on the ad hoc networking approach and the environment in which the nodes operate. For instance in ad hoc networks that consist of independent nodes and work in a hostile battlefield the physical security of single nodes may be severely threatened. Therefore in such scenarios the protection of nodes cannot rely on physical security. In contrary, in the classroom example scenario the physical security of a node is an important issue to the owner of thenode, perhaps for privacy reasons, but the breaking of the physical security does not affect the security of the system as such.

Service Principles

Ad hoc networks may apply either hierarchical or flat infrastructure both in logical and physical layers independently. As in some flat ad hoc networks the connectivity is maintained directly by the nodes themselves, the network cannot rely on any kind of centralized services. In such networks the necessary services such as the routing of packets and key management have to be distributed so that all nodes have responsibility in providing the service. As there are no dedicated server nodes, any node may be able to provide the necessary service to another. Moreover, if a tolerable amount of nodes in the ad hoc network crash or leave the network, this does not break the availability of the services. Finally, the protection of services against denial of service is in theory impossible. In ad hoc networks redundancies in the communication channels can increase the possibility that each node can receive proper routing information. Such approaches do, however, produce more overhead both in computation resources and network traffic. The redundancies in the communication paths, however, may reduce the denial of service threat and allow the system to detect malicious nodes from performing malicious actions more easily than in service provisioning approaches that rely on single paths between the source and destination.

Availability is a central issue in ad hoc networks that must operate in dynamic and unpredictable conditions. The network nodes may be idle or even be shut down once for a while. Thus the ad hoc network cannot make any assumptions about availability of specific nodes at any given time. For commercial applications using ad hoc networks availability is often the most important issue from the viewpoint of the clients. The routing protocol must guarantee the robustness of the routing fabric so that the connectivity of the network is maintained even when threatened by rapid changes in topology or attackers. Similarly, in the higher layers, the services must be able to rely on that the lower layers maintain the packet-forwarding services at any time. Finally, many ad hoc networking protocols are applied in conditions where the topology must scale up and down efficiently, e.g. due to network partitions or merges. The scalability requirements also directly affect the scalability requirements

targeted to various security services such as key management. In networks where the area of application restricts the possible size of the network, assumptions can be made about the scalability requirements of the security services as well.

Key Management Security

As in any distributed system, in ad hoc networks the security is based on the use of a proper key management system. As ad hoc networks significantly vary from each other in many respects, an environment-specific and efficient key management system is needed. To be able to protect nodes e.g. against eavesdropping by using encryption, the nodes must have made a mutual agreement on a shared secret or exchanged public keys. For very rapidly changing ad hoc networks the exchange of encryption keys may have to be addressed on- demand, thus without assumptions about a priori negotiated secrets. In less dynamic environments like in the classroom example above, the keys may be mutually agreed proactively or even configured manually (if encryption is even needed).

If public-key cryptography is applied, the whole protection mechanism relies on the security of the private key. Consequently, as the physical security of nodes may be poor, private keys have to be stored in the nodes confidentially, for instance encrypted with a system key. For dynamic ad hoc networks this is not a wanted feature and thus the security of the private key must be guaranteed with proper hardware protection (smart cards) or by distributing the key in parts to several nodes. Hardware protection is, however, never alone an adequate solution for preventing attacks as such. In ad hoc networks a centralized approach in key management may not be an available option, as there may not exist any centralized resources. Moreover, centralized approaches are vulnerable as single point of failures. The mechanical replication of the private keys or other information is an inadequate protection approach, since e.g. the private keys of the nodes simply have then a multiple possibility to be compromised. Thus a distributed approach in key management - for any cryptosystem in use - is needed, as proposed e.g. in [10].

Control in Accessing

The access control is an applicable concept also within ad hoc networking, as there usually exist a need for controlling the access to the network and to the services it provides. Moreover, as the networking approach may allow or require the forming of groups in for instance network layer, several access control mechanisms working in parallel may be needed. In the network layer the routing protocol must guarantee that no authorized nodes are allowed to join the network or a packet forwarding group such as the clusters in the hierarchical routing approach. For example in the battlefield example of the introduction the routing protocol the ad hoc network applies must control so that no hostile node can

join and leave the group undetectable from the viewpoint of the other nodes in the group. In application level the access control mechanism must guarantee that unauthorized parties cannot have accesses to services, for instance the vital key management service.

Access control is often related to the identification and authentication. The main issue in the identification and authentication is that the parties can be confirmed to be authorized to gain the access. In some systems, however, identification or authentication of nodes is not required: nodes may be given e.g. delegate certificates with which the nodes can gain access to services. In this case actual authentication mechanisms are not needed, if the nodes are able to present adequate credentials to the access control system. In some ad hoc networks services may be centralized, while in other networks they are applied in a distributed manner, which may require the use of different access control mechanisms. Moreover, the required security level in access control also affects the way the access control must be implemented. If a centralized ad hoc networking approach with low security requirements is applied - as in the classroom example - the access control can be managed by the server party with simple means such as user id - password scheme. In ad hoc networks that operate in more difficult conditions without any centralized resources as in the battlefield scenario, the implementation of access control is much more difficult. Either the access to the network, its groups and resources must be defined when the network is formed, which is very inflexible. The other possibility is to define and use a very complex, scalable and dynamic access control protocol, which brings flexibility but is prone to various kinds of attacks and it may even be impossible to apply properly and efficiently.

THREATS OF SECURITY

Types of Attacks

Attacks against ad hoc networks can be divided into two groups: Passive attacks typically involve only eavesdropping of data. Active attacks involve actions performed by adversaries, for instance the replication, modification and deletion of exchanged data. External attacks are typically active attacks that are targeted e.g. to cause congestion, propagate incorrect routing information, prevent services from working properly or shut down them completely.

External attacks can typically be prevented by using standard security mechanisms such as firewalls, encryption and so on. Internal attacks are typically more severe attacks, since malicious insider nodes already belong to the network as an authorized party and are thus protected with the security mechanisms the network and its services offer.

Thus such malicious insiders who may even operate in a group may use the standard security means to

actually protect their attacks. These kind of malicious parties are called compromised nodes, as their actions compromise the security of the whole ad hoc network.

(DoS) Denial of Service

Recent wireless research indicates that the wireless MANET presents a larger security problem than conventional wired and wireless networks[3]. Distributed Denial of Service (DDoS) attacks has also become a problem for users of computer systems connected to the Internet. A DDoS attack is a distributed, large-scale attempt by malicious users to flood the victim network with an enormous number of packets. This exhausts the victim network of resources such as bandwidth, computing power, etc. The victim is unable to provide services to its legitimate clients and network performance is greatly deteriorated.

A denial of service (DoS) attack is characterized by an explicit attempt by an attacker to prevent legitimate users of a service from using the desired resources [6]. Examples of denial of service attacks include

- attempts to "flood" a network, thereby preventing legitimate network traffic
- attempts to disrupt connections between two machines, thereby preventing access to a service
- attempts to prevent a particular individual from accessing a service
- attempts to disrupt service to a specific system or person

The denial of service attack has many forms: the classical way is to flood any centralized resource so that it no longer operates correctly or crashes, but in ad hoc networks this may not be an applicable approach due to the distribution of responsibility. Distributed denial of service attack is a more severe threat: if the attackers have enough computing power and bandwidth to operate with, smaller ad hoc networks can be crashed or congested rather easily.

Disclosure

Any communication must be protected from eavesdropping, whenever confidential information is exchanged. Also critical data the nodes store must be protected from unauthorized access. In ad hoc networks such information can include almost anything e.g.specific status details of a node, the location of nodes, private or secret keys, passwords and -phrases and so on. Sometimes the control data is more critical information in respect of the security than the actual exchanged data. For instance the routing directives in packet headers such as the identity or location of the nodes can sometimes be more valuable than the application-level messages. This applies especially in critical military applications. For instance in the battlefield scenario the data of a "hello" packet exchanged between nodes may not be

as interesting from the viewpoint of the enemy. Instead the identities of the observed nodes - compared to the previous traffic patterns of the same nodes - or the detected radio transmissions the nodes generate may be the information just the enemy needs to launch a well-targeted attack. On the contrary, in the classroom example the disclosure of exchanged or stored information is critical "only" from the viewpoint of a person's privacy.

TECHNIQUES USED FOR SECURITY

DDM

Dynamic Destination Multicast protocol (DDM) is a multicast protocol that is relatively different from many other multicast-based ad hoc protocols. In DDM the group membership is not restricted in a distributed manner, as only the sender of the data is given the authority to control to which the information is really delivered. In this way the DDM nodes are aware of the membership of groups of nodes by inspecting the protocol headers.

The DDM approach also prevents outsider nodes from joining the groups arbitrarily. This is not supported in many other protocols directly; if the group membership and the distribution of source data have to be restricted, external means such as the distribution of keys have to be applied.

DDM has two modes of operation: the stateless mode and the soft-state mode. In the stateless mode the maintenance of multicast associations and restriction of group membership are handled totally by encoding the forwarding information in a special header of the data packets; the nodes do not have to store state information. This kind of reactiveapproach thus guarantees that there are no vainless exchange of control data during idle periods. Thus in small ad hoc networks that need not scale up substantially, this kind of ultra-reactive approach can be extremely useful. The soft-state mode, on the other hand, requires that the nodes remember the next hops of every destination and thus need not fill up the protocol headers with every destination. In both modes the nodes must always be able to keep track of the membership of the groups. According to the authors, DDM is best suited for dynamic networks having small multicast groups. Currently the DDM draft ([8]) does not, however, propose any solutions for securing the DDM networks as such. Moreover, it does not provide any suggestions for a concrete protocol that handles the necessary access control needed in the restriction of group membership.

OLSR

Optimized Link State Routing protocol (OLSR), as defined in [7], is a proactive and table driven protocol that applies a multi-tiered approach with multi-point relays (MPR). MPRs allow the network to apply scoped flooding, instead of full node-to-node flooding, with which the amount of exchanged

control data can substantially be minimized. This is achieved by propagating the link state information about only the chosen MPR nodes.

Since the MPR approach is most suitable for large and dense ad hoc networks, in which the traffic is random and sporadic, also the OLSR protocol as such works best in these kind of environments. The MPRs are chosen so that only nodes with one-hop symmetric (bi- directional) link to another node can provide the services. Thus in very dynamic networks where there exists constantly a substantial amount of uni-directional links this approach may not work properly. OLSR works in a totally distributed manner, e.g. the MPR approach does not require the use of centralized resources. The OLSR protocol specification does not include any actual suggestions for the preferred security architecture to be applied with the protocol. The protocol is, however, adaptable to protocols such as the Internet MANET Encapsulation Protocol (IMEP), as it has been designed to work totally independently of other protocols.

ODMRP

On-Demand Multicast Routing Protocol (ODMRP) is a mesh-based multicast routing protocol for ad hoc networks, specified in [10]. It applies the scoped flooding approach, in which a subset of nodes - a forwarding group - may forward packets. The membership in the forwarding groups are built and maintained dynamically on-demand. The protocol does not apply source routing. ODMRP is best suited for MANETs where the topology of the network changes rapidly and resources are constrained. ODMRP assumes bi-directional links, which somewhat restricts the potential area of application for this proposal; ODMRP may not be suitable for use in dynamic networks in which nodes may move rapidly and unpredictably and have varying radio transmission power. Currently ODMRP does not define or apply any security means as such, "the work is in progress". The forwarding group membership is controlled with the protocol itself, though.

AODV and MAODV

Ad Hoc On-Demand Distance-Vector routing protocol (AODV), defined in [5], is an unicast- based reactive routing protocol for mobile nodes in ad hoc networks. It enables multi-hop routing and the nodes in the network maintain the topology dynamically only when there is traffic. Currently AODV does not define any security mechanisms whatsoever.

The authors identify the necessity of having proper confidentiality and authentication services within the routing, but suggest no solutions for them. The IPSec is, however, mentioned as one possible solution. Multicast Ad Hoc On-Demand Distance-Vector routing protocol (MAODV), specified in

[16], extends the AODV protocol with multicast features. The security aspects currently noted in the design of MAODV are similar to the AODV protocol.

TBRPF

Topology Broadcast based on Reverse-Path Forwarding (TBRPF), as defined in [2], is a pure proactive, link-state routing protocol for the ad hoc networks that can also be applied as the proactive part in hybrid solutions. Each of the nodes of the network in TBRPF carry state information of each link of the network, but the information propagation is optimized by applying reverse-path forwarding instead of the costly full flooding or broadcast techniques. TBRPF operates over IPv4 in ad hoc networks and can also be applied within hierarchical network architecture. The authors of the proposal, however, do not suggest any specific mechanisms for securing the protocol. Finally, the protocol, just as every other ad hoc network routing protocol, can be protected with IPSec, but this approach is not currently officially in use within TBRPF.

CONCLUSION AND FUTURE SCOPE

Security is an important topic that needs to be addressed when designing networks in MANET environments. This topic involves far more than network routing protocols. In addition, existing security methods are insufficient. They are not geared towards the specialized needs of a MANET. The areas of concern within MANET data communication are raised. Future research will need to begin to resolve DoS attack very carefully because it is one of the most harmful attacks for the network. Along with these issues, standardized benchmarks and criteria for Evaluation must be established so that proposed protocols and methods can be legitimately compared.

REFERENCES

[1] Hao Yang, haiyun luo, Fan Ye, Songwu Lu and Lixia Zhanng, "Security on Mobile Ad Hoc Networks: Challenges and Solutions" 1536-1284/04/IEEE Wireless Communications Feb. 2004

[2] S.Buchegger and J.L. Boudec, "Performance Analysis of the Confident Protocol in dynamic Ad Hoc Networks IEEE/ACM Symp., 2002

[3] S. Corson and J.Macker, "Mobile Ad Hoc Networking routing Protocol Performance Issues and evaluation considerations", RFC2501, Jan. 1999

[4] Corson, M., Freebergyser, J., and Sastry, A., "Mobile Ad Hoc Networking: Editorial, "Mobile Networks and Applications, 4(3): pp. 137-138, 1999

[5] Singh, S., Woo, M., and Raghavendra, C. Power Aware Routing in Mobile Ad Hoc Networks. In Proc. 4th International Conf. on Mobile Computing and Networking (MOBICOM '98), pp. 181-190, October, 1998.

[6] Guo, Y., Pinotti, M., and Das, S., "A New Hybrid Broadcast Scheduling Algorithm for Assumetric Communication Systems," ACM Mobile Computing and Communications Review, 5(4): pp. 39-54, 2001

[7] Claude Castelluccia, Nitesh Saxena, Jeong Hyun Yi, "Robust self-keying mobile adhoc networks" Computer Networks 51 (2007) 1169–1182 1389-1286 2006 Elsevier B.V. doi:10.1016/j.comnet.2006.07.009

[8] C.Zhuand M.Corson, QoS routing for mobile adhoc networks, tech. report, CSHCN Technical Report 2001.

[9] Kärpijoki, V. Signalling and Routing Security in Mobile Ad Hoc Networks. Proceedings of the Helsinki University of Technology, Seminar on Internetworking - Ad Hoc Networks, Spring 2000. [referred 25.4.2000]

[10] Lee, S.-J. et al. On-Demand Multicast Routing Protocol (ODMRP). IETF draft, January 2000 (expired).

Security Analysis of Cloud Computing

Anju Chhibber[1], Dr. Sunil Batra[2]

[1]Lecturer, Department of Computer Science & Applications, Guru Nanak Khalsa Inst. of Tech. & Management Studies (GNKITMS), Yamuna Nagar, India

[2]Assistant Professor, Department of Computer Applications, Chandigarh Group of Colleges Landran, Mohali (Pb), India

ABSTRACT

Cloud computing is a model that uses the concept of utility computing that gives on-demand services to the end users. Cloud computing has the flexibility to produce shared resources over the internet and avoid serious installation price for it. However in conjunction with those benefits there's additionally a chance wherever a malicious user can infiltrate the cloud by impersonating a legitimate user that affects many shoppers who are sharing the cloud. This paper explore the cloud security problems faced by cloud service consumer such as data, privacy, and infected application and security problems and discuss some remedial measure to scale back the security risk.

1INTRODUCTION

'Cloud computing' is an emerging information technology for storing, processing and use of data from remotely located computers that can be accessed over the internet. This provides unlimited computing power on demand to users, which does not require major capital investments to fulfill their needs. In addition to that with the help of an internet connection they can retrieve their data from anywhere.

There are three types of cloud which includes private, public and hybrid cloud. The private cloud is owned by a single organization and public clouds are shared on a larger scale. Private cloud provides better control and more flexibility. Private Cloud and Public Cloud jointly makes Hybrid cloud, and this cloud is used by most of the industries. The advantages of cloud computing are very appealing but nothing is ideal. Cloud got several problems once it involves security particularly on data thievery, data loss and Privacy. This paper explores the cloud security threats and discusses some solutions to handle the security issue.

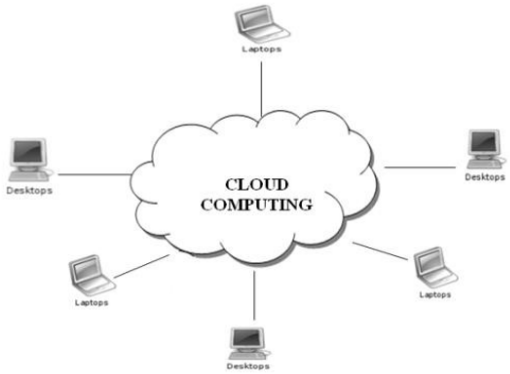
2. CLOUD COMPUTING SECURITY THREATS

The biggest challenge in implementing successful Cloud computing technologies is managing the security. As the sensitive applications and data are moved into the cloud data centers, run on virtual computing resources in the form of virtual machine which may cause many security concerns. Top seven security threats to cloud computing that are discovered by "Cloud Security Alliance" (CSA) are

a) Nefarious Use of Cloud Computing: It is the top threat identified by the CSA. In this approach

attackers can penetrate a public cloud to upload malware to thousands of computers and use the power of the cloud infrastructure to attack other machines.

b) Insecure API (Application Programming Interfaces): APIs are set of software interfaces that are used by customers to interact with cloud services. When third parties begin to create that application then user take risks with the supply, confidentiality and integrity of data.



c) Shared Technology Vulnerabilities: As cloud provider platform being shared by different user there may be possibility that information belonging to different customers reside on same data center. Therefore Information leakage may arise as by mistake information for one customer is given to other.

d) Data Loss/Leakage: Data loss is a common problem in cloud computing. If the cloud computing service provider close up his services due some financial or legal problem then there will be a loss of data for the user.

e) Traffic Hijacking: Traffic hijacking is another issue that cloud users need to be aware of. These threats include man-in-the-middle attacks, spam campaigns and denial-of service attacks.

f) Malicious insiders: Such threats include fraud, damage and theft or loss of confidential information caused by trusted insiders. The malicious insiders can have the ability to infiltrate organizations and assets like productivity losses, brand damage and financial impact.

3. EXISTING SOLUTIONS FOR SECURITY THREATS

a) Mirage Image Management System: The integrity of VM images are the foundation for the overall security of the cloud. In this system use of Filters alleviate the risk in an efficient way. This system stores all the revisions which allow the user to go back to the previous version. The default access permission for an image is private so that only owner and system administrator can access the image and hence untrusted parties cannot access the image.

b) Client Based Privacy Manager: client based privacy manager helps to reduce the risk of data leakage and loss of privacy of the sensitive data processed in the cloud, and provides additional privacy related benefits.

c) Transparent Cloud Protection System (TCPS): It is a protection system which is intended to protect the integrity of guest Virtual Machines (VM) by allowing the host to monitor guest VMs and effective in detecting most kind of attacks. The system can detect an intrusion try over a guest and, if needed by the safety policy, takes applicable actions against the attempt or against the compromised guest.

4. OTHER EXISTING SOLUTIONS TO MANAGING CLOUD COMPUTING SECURITY

Managing and controlling Cloud issues will need to address but not limited to the following:

a) Cloud Governance: Cloud computing policies and procedures should be put in place to protect the cloud from potential of threats, hacks and the loss of information. The protection of data in the cloud is a key consumer concern particularly for committing fraudulent activities and financial exploitation. With governance and security in place, Cloud computing can be used safely and with confidence.

b) Cloud Transparency: Transparent security will make compulsory for cloud providers to disclose adequate information about their security policies and practices. SLA is one of the most significant protocols to ensure transparency within Cloud computing environment. The SLA is the only legal agreement between the service provider and client which includes the following rules:

i. Services to be delivered, performance,
ii. Tracking and Reporting
iii. Legal Compliance
iv. Security responsibility

c) Cloud Computing Security Impact: As computer makers, employers and universities install cloud based tools on desktops, several users could fail to understand that they're actually victimisation an online based service. A HTTPS encrypted connection takes significantly more processing power and memory for a Web server to provide than a normal web connection.

WS-Security assists with SOAP messages by shaping the header that carries the WS-Security extensions. The cloud computing moves a lot of of a user's traditional activity to the online browser.

data in a single place. as such it's doable for malicious websites to take advantage of browser to steal data related to different existing or previous browsing sessions, like a logged in email account or on-line banking session. it's for this reason that some security specialists suggest that customers use one web browser for general surfing, and another for additional sensitive tasks, like online banking. Potential Cloud organization sought to remember that it's going to become easier for attackers to threaten clouds by moving towards one cloud interface.

5. CONCLUSION

Although Cloud computing can be seen as a new technology which revolutionize the way of using the Internet. However one must be very careful to understand the limitations and security risks posed in utilizing these technologies. In this paper key security challenges are highlighted which are currently faced by cloud industry.

BIBLIOGRAPHY

1. *Dawei Sun, G. C. (2011). Surveying and Analyzing Security, Privacy and Trust Issues. Advanced in Control Engineering and Information Science , 2852 – 2856.*
2. *Dimitrios Zissis, D. L. (2012). Addressing cloud computing security issues. Future Generation Computer Systems , 28 (3), 583-592.*
3. *Kshetri, N. (4 July 2012). Privacy and security issues in cloud computing: The role of institutions and institutional evolution. Telecommunications Policy , [In Press].*
4. *Latifa Ben Arfa Rabai, M. J. (2012). A cybersecurity model in cloud computing environments. Journal of King Saud University – Computer and Information Sciences (25), 63-75.*
5. *Ramgovind S, E. M. (n.d.). The Management of Security in Cloud Computing.*
6. *Shilpashree Srinivasamurthy, D. Q. (n.d.). Survey on Cloud Computing Security.*
7. *Winkler, V. (. (2011). Introduction to Cloud Computing and Security. Cloud Computer Security Techniques and Tactics.*

A Review of Distributed Shared Memory

Pankaj Sharma, Naveen Malik, Naeem Akhtar,
Rahul, Hardeep Rohilla
Student, CSE,
Dronachraya College of Engineering, Gurgaon

ABSTRACT

Distributed shared memory (DSM) systems have attracted considerable research efforts recently, since they combine the advantages of two different computer classes: shared memory multiprocessors and distributed systems. The most important one is the use of shared memory programming paradigm on physically distributed memories. In the first part of this paper, one possible classification taxonomy, which includes two basic criteria and a number of related characteristic, is proposed and described. According to the basic classification criteria-implementation level of DSM mechanism--systems are organized into three groups: hardware, software, and hybrid DSM implementations. The second part of the paper represents an almost exhaustive survey of the existing solutions in an uniform manner, presenting their DSM mechanisms and issues of importance for various DSM systems and approaches.

Software DSM systems can be implemented in an operating system (OS), or as a programming library and can be thought of as extensions of the underlying virtual memory architecture. When implemented in the OS, such systems are transparent to the developer; which means that the underlying distributed memory is completely hidden from the users.

HISTORY:-

Memory mapped files started in the MULTICS operating system in the 1960s. One of the first DSM implementations was Apollo. One of the first system to use Apollo was Integrated shared Virtual memory at Yale (IVY). DSM developed in parallel with shared-memory multiprocessors.

INTRODUCTION:-

Distributed Shared Memory (DSM), in Computer Architecture is a form of memory architecture where the (physically separate) memories can be addressed as one (logically shared) address space. Here, the term shared does not mean that there is a single centralized memory but shared essentially means that the address space is shared (same physical address on two processors refers to the same location in memory). Distributed Global Address Space (DGAS), is a similar term for a wide class of software and hardware implementations, in which each node of a cluster has access to shared memory in addition to each node's non-shared private memory.

Software DSM systems can be implemented in an operating system (OS), or as a programming library

and can be thought of as extensions of the underlying virtual memory architecture. When implemented in the OS, such systems are transparent to the developer; which means that the underlying distributed memory is completely hidden from the users. In contrast, Software DSM systems implemented at the library or language level are not transparent and developers usually have to program differently. However, these systems offer a more portable approach to DSM system implementation.

Software DSM systems also have the flexibility to organize the shared memory region in different ways. The page based approach organizes shared memory into pages of fixed size. In contrast, the object based approach organizes the shared memory region as an abstract space for storing shareable objects of variable sizes. Another commonly seen implementation uses a tuple space, in which the unit of sharing is a tuple.

Shared memory architecture may involve separating memory into shared parts distributed amongst nodes and main memory; or distributing all memory between nodes. A coherence protocol, chosen in accordance with a consistency model, maintains memory coherence Distributed shared memory (DSM) is an abstraction used for sharing data between computers that do not share physical memory. Processes access DSM by reads and updates to what appears to be ordinary memory within their address space. However, an underlying runtime system ensures transparently that processes executing at different computers observe the updates made by one another. It is as though the processes access a single shared memory, but in fact the physical memory is distributed. The main point of DSM is that it spares the programmer the concerns of message passing when writing applications that might otherwise have to use it. DSM is primarily a tool for parallel applications or for any distributed application or group of applications in which individual shared data items can be accessed directly. DSM is in general less appropriate in client-server systems, where clients normally view server-held resources as abstract data and access them by request (for reasons of modularity and protection). However, servers can provide DSM that is shared between clients. For example, memory-mapped files that are shared and for which some degree of consistency is maintained are a form of DSM. (Mapped files were introduced with the MULTICS operating system [Organick 1972].) Message passing cannot be avoided altogether in a distributed system: in the absence of physically shared memory, the DSM runtime support has to send updates in messages between computers. DSM systems manage replicated data: each computer has a local copy of recently accessed data items stored in DSM, for speed of access. The problems of implementing DSM are related to those discussed in Chapter 15, as well as those of caching shared files discussed in Chapter 8.

One of the first notable examples of a DSM implementation was the Apollo Domain file system. 1983],

in which processes hosted by different workstations share files by mapping them simultaneously into their address spaces. This example shows that distributed shared memory can be persistent. That is, it may outlast the execution of any process or group of processes that accesses it and be shared by different groups of processes over time. Examples of such systems include:

- Kerrighed
- OpenSSI
- MOSIX
- TreadMarks

.

1. DSMClassification:-

In order to provide a wide and extensive overview in the field of DSM, possible platforms for classification and a set of relevant parameters that must be considered in DSM design are proposed. The selection of classification criteria can be taken conditionally, since some of the parameters could also be adopted as the platform for classification. Our choice of classification criteria relies on the possibility to classify all existing systems into the appropriate non-overlapping subsets of systems with common general advantages and drawbacks.

The criterion: DSM implementation level

Types:

1. Hardware
2. Software

2.1. Operating system

2.1.1. Inside the kernel

2.1.2. Outside the kernel

2.2. Runtirne library routines

2.3. Compiler-inserted primitives

3. Hardware/software combination

The level of DSM implementation affects both the programming model and the overall system performance. While the hardware solutions bring the total transparency to the programmer, and achieve very low access latencies, software solutions can better exploit the application behavior and represent the ideal polygon to experiment with new concepts and algorithms. As the consequence, the number of software DSM systems presented in the open literature is considerably higher, but the systems intending to become commercial products and standards are mostly hardware-oriented.

Parameters closely related to the DSM implementation level

Some important characteristics of the system are often (but not necessarily) closely related, or even determined by this criterion.

Architectural configuration of the system affects the system performance, since it can offer or restrict a good potential for parallel processing of requests related to the DSM management. It also strongly affects the scalability. Since a system applying a DSM mechanism is usually organized as a set of clusters connected by an interconnection network, architectural parameters include:

a) Cluster configuration (single/multiple processors, with/without, shared&ivate, single/multiple level caches, etc.)

b) Interconnection network (bus hierarchy, ring, mesh, hypercube, specific LAN, etc.)

Cluster configuration is usually very important for the hardware-oriented proposals that integrate the mechanisms of cache coherence on the lower level with the DSM mechanisms on the higher level of the system organization, or even store all shared data in large caches. Cluster configuration is mostly transparent for software solutions, It includes the memory organization and the placement of directory, as well. Almost all types of interconnection networks found in multiprocessors and distributed systems have also been used in DSM systems, The majority of software-oriented DSM systems were actually build on the top of Ethernet, although some of the solutions tend to be architecture independent and portable to various platforms. On the other hand, topologies such as bus hierarchy or mesh are typical for hardware solutions. The choice of topology can be also very important for the implementation of DSM algorithm, since it affects the possibility and cost of broadcast and multicast transactions. Shared data organization represents the global layout of shared address space, as well as the size and organization of data items in it, and can be distinguished as:

a) Structure of shared data (non structured or structured into objects, language types. etc.)

b) Granularity of coherence unit (word, cache block, page, complex data structure, etc.)

The impact of this organization to the overall system performance is closely related to the locality of data access.

Hardware solutions always &al with non-structured data objects (typically cache blocks), while many software implementations tend to use data items that represent logical entities, in order to take advantage of the locality naturally expressed by the application. On the other hand, some software solutions, based on virtual memory mechanisms, organize data in larger physical blocks (pages), counting on the coarse-grain sharing.

The second criterion: DSM algorithm

Types:

1. SRSW (Single Reader/Single Writer)

1.1. Without migration

1.2. With migration

2. MRSW (Multiple Reader/Single Writer)

3. MRAM' (Multiple Reader Multiple Writer)

This classification is based on the possible existence of multiple copies of a data item, also considering access rights of those copies. The complexity of coherence maintenance is strongly dependent on the introduced classes. In order to explore the properties of application behavior, including typical read/write patterns, while keeping the acceptable complexity of the algorithm, many solutions were proposed, among which h4RSW algorithms represent the majority.

Parameters closely retied to the DSM algorithm

a) Responsibility for the DSMmanagement (centralized, distributed &wed, distributed/dynamic)

b) Consistency model (strict, sequential, processor, weak ,release, lazy release, entry, etc.)

c) Coherence policy (write-invalidate, write-update, type specific, etc.)

Responsibility for DSM management can be centralized or distributed, and it determines which site has to handle actions related to the consistency maintenance in the system. Centralized management is easier to implement, but suffers from the lack of fault tolerance, while the distributed management can be defined statically or dynamically, eliminating bottlenecks and providing scalability. Distribution of responsibility for DSM management is closely related to the distribution of directory information, that can be organized in the form of linked lists or trees. Memory consistency model defines the legal ordering of memory references issued by some processor and observed by other processors. Stronger forms of consistency typically increase the memory access latency and the bandwidth requirements, and simplify the programming. More relaxed models result in better performance, at the expense of a higher involvement of the programmer in synchronizing the accesses to shared data. In strive to achieve an optimal behavior, systems with multiple consistency models adaptively applied to appropriate data types have been recently proposed. Coherence policy determines whether the existing copies of the data item being written to at one site will be immediately updated, or just invalidated on the other sites. The choice of coherence policy is related to the granularity of shared data. For very fine grain data items, the cost of update message is approximately the same as the cost of invalidation message.

Therefore, update policy is typical for systems with word-based coherence maintenance, and invalidation is used in coarse-grain systems. The efficiency of an invalidation approach is increased when the sequences of read and write to the same data item by various processors are not highly interleaved.

2. DSM IMPLEMENTATION:-

2.1. Hardware Level DSM Implementations:-

Hardware implementations of the DSM concept usually extend the principles found in traditional cache coherence schemes of scalable shared-memory architectures, Therefore, the unit of sharing is smaller-typically cache line size. Communication latency is considerably reduced, based on the advantage of megraing sharing, that also minimizes the effects of false sharing and thrashing. Searching and directory functions are much faster, compared to the software level implementations, as well. According to the general properties of memory system architecture, three groups of hardware DSM systems are regarded as especially interesting:

- CC-NUMA (Cache Coherent Non-Uniform Memory Architecture)
- COMA (Cache-Only Memory Architecture)
- RMS (Reflective Memory System architecture)

In a CC-NUMA system, the shared virtual address space is statically distributed across local memory modules of clusters. It is accessible by the local processors and by processors from other clusters, with quite different access latencies. The DSM mechanism relies on directories with organizations varying from a full-map storage to different dynamic organizations, such as single or double linked lists and fat trees. In order to minimize latency, static partitioning of data should be done with extreme care, in order to maximize the frequency of local access. The invalidation mechanism is typically applied, while some relaxed memory consistency model can be used as a source of performance improvement. Typical representatives of this type of DSM approach are Dash and SCI.

COMA architecture provides the dynamic partitioning of data in the form of distributed memories, organized as large second-level caches (attraction memories). There is no physical memory home location predetermined for particular data item, which can be simultaneously replicated in multiple caches. The existing COMA architectures are characterized by hierarchical network topologies that simplify two main problems in thistype of systems: finding an item and replacement of a cache block. In COMA architectures, the distribution of data across attraction memories is dynamically adaptable to the application behavior, therefore, they are less sensitive to static distribution of data than the NUMA architectures. Increased storage overhead for keeping the information typical for cache memory is

inherent to the COMA architecture. However, some findings pointed out that this approach means an acceptably low amount of the overall system memory. Two most relevant representatives of COMA systems are DDM and KSR 1.Reflective memory systems are characterized with hardware-implemented update mechanism on the low level of data granularity. Some parts of local memory in each cluster can be declared as shared, and appropriately mapped into the common virtual space. Coherence maintenance of shared regions in these systems is based on full-replication algorithm. Following the assumption that all data written will be soon read by other sharing processors, those systems immediately propagates very change of all sharing sites, using a broadcast or multicast mechanism. Because of the property of "reflection," this kind of memory is also called "mirror memory." It results in high cost of write operations, especially when multiple writes to the same location occur, consequently, this architecture is the most convenient for the applications with a lower write frequency.

2.2. Software level DSM Implementations:-

The basic principle of the first software-implemented DSM mechanisms was quite alike to that of the virtual memory management, except that on page fault the data are supplied from local memory of the remote cluster instead from the local secondary storage. Software implementations of the DSM concept are usually built as a separate layer on the top of message passing model. According to the implementation level, several types of software- oriented DSM approaches can be recognized.

1. Compiler implementations. In the cases where the DSM paradigm is applied at the level of parallel programming language, the shared address space is usually structured into logical units of sharing Therefore, shared data have to be declared as a specific type in the source program. In this approach, a can accesses to shared data are automatically converted into synchronization and coherence primitives. Language implementation can be portable between various systems, and recompiled if appropriate run-time primitives are available. Linda is an example of this software approach to DSM implementation.

2. User-level runtime packages. The DSM mechanism is implemented by virtue of the run- time library routines, which are to be linked with an application program that uses the shared virtual address space. This approach is not only convenient for experimenting, but also flexible and easy to implement. IVY and Mermaid systems are based on run-time library routines.

3.1. Operating system level &side the kernel). The interaction of scheduling, interrupt processing, and the application behavior can be efficiently examined if the DSM model is incorporated into the operating system kernel. The advantage of this approach is that the semantics of the underlying OS

architecture can be preserved; hence, the applications can be ported from the local environment to the distributed system without being recompiled. Mirage is an example of system built according to this method.

3.2. Operating system level (outside the kernel). The DSM mechanism is incorporated in the specialized software controller, that can be (with minor modifications) used by different kernels. The same DSM mechanism can be used both by user and by the operating system kernel objects. One of the systems that follow this implementation level is Clouds.

The above classification should be taken conditionally, since all programming language implementations require some operating system support. Also, some programmers hints, at the level of the language, can help the run-time implementation to become more efficient. Software implementations are clearly inferior in performance to hardware implementations, but they are less expensive, can be suitable for a variety of underlying architectures, and can better take the advantage of the application characteristics. Problem- oriented shared memory, DSM in heterogeneous environments, and various sophisticated consistency mechanisms are mostly implemented in software.

2.3. Hybrid DSM Implementations:-

The integration of software and hardware methods, competitive management of DSM, and various consistency models, seem to be the most promising approach in the future of DSM. The idea to implement a combination of hardware and soft& ware is explored in the efforts to achieve scalability, limited in directory based schemes, that use the full-map hardware directories. Based on the observation that only a few simultaneously shared copies of the same shared data exist on average, the solution was found in which only a limited number of pointers for each directory entry are implemented in hardware. When more directory storage is needed, it is managed by software. This principle is implemented in the MIT Alewife system. In order to gain better performance of DSM systems, researchers experiment lately with the use of multiple protocols within the same system, and even integrate message passing with the DSM mechanism. Innovative consistency models are also being implemented, requiring additional activities of the programmer to suit the needs of application. In order to handle complexity of those, basically software solutions, special programmable protocol processors are added to the system, as it was done in the Stanford FLASH system.

3. CONCLUSION

The intention of this survey was to provide an extensive coverage of a11 relevant topics in an increasingly important area -distributed shared memory computing. A special attempt has been made to

give the broadest overview of the proposed and existing approaches, in a uniform organizational manner. Because of the combined advantages of the shared memory and distributed systems, DSM solutions appear to be the most appropriate way toward large-scale high-performance systems with a reduced cost of parallel software development. In spite of that, building of successful commercial systems that follows the DSM paradigm is still in its infancy; consequently experimental and research efforts prevail. Therefore, the DSM field remains a very active research area. Some of the promising research typical applications and system configurations, synergistic combining of hardware and software implementations of the DSM concept, integrating of the shared memory and message passing programming paradigms, innovative system architecture (especially memory system), comparative performance evaluation, etc. From this point of view, tirther investment in exploring, developing, and implementing DSM systems seems to be quite justified and promising.

REFERENCE:-

1. *http://en.wikipedia.org/wiki/Distributed_shared_memory*
2. *http://books.google.co.in/books/about/Distributed_Shared_Memory.html?id=YgXPsumHu4C&rediresc=y*
3. *http://www.studymode.com/*
4. *[Bershad1993] - Bemhad, B., N, Zekauskas M., J., Sawdon, W., A, "The Midway Diszibuted Shared Memory System," COMPCON 93. February 1993,*
5. *[Li1989,] - Li, K. Shared Virtual Memory on Loosely Coupled Multiprocessors. PhD thesis, Department of Computer Science, Yale University, September 1986.*
6. *[Tanenbaum1992] - Bal, H., E., Tanenbaum, A, S., "Distributed programming with shared data," International Conjerance on Computer Languages '88, October 1988.*
7. *[Kulkarni et al. 1993] - Kulkarni, D.~C. et~al. Structuring Distributed Shared Memory with the "pi" Architecture. In Proc. of the 13th Int'l Conf. on Distributed Computing Systems (ICDCS-13).*
8. *[AHUJA86] - Ahuja, S., Carriero, N., Gelernter, D.,"Linda and Friends," IEEE Computer, Vol. 19, No. 8. May 1986,*
9. *[AGARW90] - Agarwal, A., Lim, B., Kranq D., Kubiatowin, J., "APRIL: A Processor Architecture for Multiprocessing," Proceedings of the 17th Annual International- Symposium on Com$e;. Architecture, 1990,*

Instructions for Authors

Essentials for Publishing in this Journal

1 Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.

2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.

3 All our articles are refereed through a double-blind process.

4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

All articles for this journal must be submitted using our online submissions system. http://enrichedpub.com/ . Please use the Submit Your Article link in the Author Service area.

Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd**. The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page.

Type of Articles

Classification of articles is a duty of the editorial staff and is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).

2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);

3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);

4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);
2. Informative contribution (editorial, commentary, etc.);
3. Review (of a book, software, case study, scientific event, etc.)

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250-Word abstract should be placed between the title and the keywords with the body text to follow. Besides an abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

Keywords

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

Acknowledgements

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

Tables and Illustrations

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the text and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

Citation in the Text

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.
The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.

Address of the Editorial Office:

Enriched Publications Pvt. Ltd.
S-9,IInd FLOOR, MLU POCKET,
MANISH ABHINAV PLAZA-II, ABOVE FEDERAL BANK,
PLOT NO-5, SECTOR -5, DWARKA, NEW DELHI, INDIA-110075,
PHONE: - + (91)-(11)-45525005