IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 09, Issue 5 (May. 2019), ||S (XV) || PP 55-61

"Secured Data Retrieval Disruption Tolerent Network"

Pushpraj Deshkar[1], Mr. Amit Pampatwar[2], Ms. Raana Syeda[3]
Department of Computer Science & Engineering, Jhulelal Institute of Technology, Nagpur, India
e-mail: pushprajdeshkar68@gmail.com, amitascent@gmail.com, raana,syeda@jit.org.in

Abstract: *The information leak of sensitive data on systems has a serious threat to organization data security. Statistics show that the improper encryption on files and communications due to human errors is one of the leading causes of information loss. So there a need tools to identify the exposure of sensitive data by monitoring the content in storage and transmission. However, detecting the exposure of sensitive data information is challenging due to data transformation in the content. Transformations result in highly unpredictable leak patterns. In this paper, the sequence alignment method used for detecting data leak patterns. This algorithm is designed for detecting long and inexact sensitive data patterns. This detection is paired with a comparable sampling algorithm, which allows one to compare the similarity of two separately sampled sequences. The method have efficient detection accuracy in recognizing pattern data leaks. It implement a the algorithms in data processing to get high analysis data. In the case of collective privacy preservation, organizations have to cope with some interesting conflicts. The personal information verifies by analysis processes that give new facts about users' shopping patterns, hobbies, or preferences, these facts could be used in recommender systems to predict or affect their future shopping patterns. In general, this scenario is beneficial to both users and organizations. However, when organizations share data in a collaborative project, the goal is not only to protect personally identifiable information but also sensitive knowledge represented by some strategic patterns. To demonstrate the high multithreading scalability of the data leak detection method required by a requirement of organization.*

Keywords: *Information leak detection, content inspection, sampling, alignment, dynamic programming, sensitive data patterns.*

I. Introduction

To minimize the exposure of sensitive data and documents, an organization needs to prevent clear text sensitive data from appearing in the storage or communication. In today's increasingly digital world, there is often a tension between safeguarding privacy and sharing information. Although, in general, sensitive data clearly needs to be kept confidential data owners are often motivated, or forced, to share sensitive information Privacy-Preserving Sharing of Sensitive Information , and proposes one efficient and secure instantiation that functions as a privacy shield to protect parties from disclosing more than the required minimum of sensitive information. We model in the context of simple database-querying applications with two parties: a server that has a database, and a client, performing simple disjunctive equality queries Detecting the exposure of sensitive information is challenging due to data transformation in the content. Transformations result in highly unpredictable leak patterns. In this paper, we utilize sequence alignment techniques for detecting complex data-leak asymmetric cryptography, facilitate the creation of a verifiable association between a public key and the identity other attributes of the holder of the corresponding private key, for uses such as authenticating the identity of a specific entity, ensuring the integrity of information, providing support for non repudiation, and establishing an encrypted communications section.

II. Brief Literature Survey

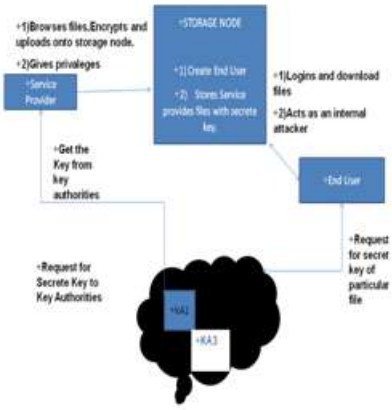
Big data analysis system concept for detecting unknown attacks:Sung-Hwan Chung. 16-19 Feb. 2014(IEEE). Unknown cyber-attacks are increasing because existing security systems are not able to detect them. Big data analysis techniques that can extract information from a variety of sources to detect future attacks. The event of new and previously unknown attacks, detection rate becomes very low and false negative increases. To defend against these unknown attacks. Does not detect future Advanced Persistent Threat (APT) detection.[1]

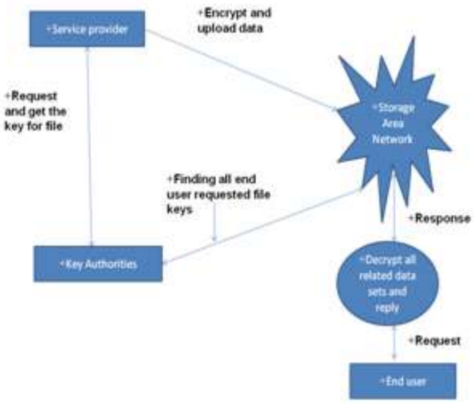
Big Data Analytics with Hadoop to analyze Targeted Attacks on Enterprise Data: Bhawna Gupta, Dr.Kiran Joyti in Journal of Computer Science and Information Technologies, Vol.5, 2014, (IEEE). Big data security analytics is used for the growing practice of organization to gather and analyze security data to detect vulnerabilities and intrusions. Security and Information Event Monitoring (SIEM) system. The malicious and targeted attacks have become main subject for government, organization or indust. Big data analytics is the process of analyzing big data to find hidden patterns, unknown correlations and other useful information that can be extracted to make better decisions. It is used effectively and at the same time, hackers can leave their targets forever.[2] Zero Day Attack Signatures Detection Using Honeypot: IEEE 29-31 May 2013. Unexpected behavior. Fault distribution studies show that there is a correlation between the number of lines of code and the number of faults. LCS algorithm for the data verification on the packet content. When a Zero-day attack is a computer threat that tries to exploit computer application vulnerabilities that are unknown to others or undisclosed to the software developer. Vulnerability window which is the time between the first exploitation of vulnerability and when software developers start to develop a countermeasure to that threat. [3]

Cloud Model based Outlier Detection Algorithm for Categorical Data: Dajiang Lei Liping Zhang And Lisheng Zhang, Vol. 6, No. 4, August, 2013. Numerical data but there will be a large number of categorical data in real life. Some outlier detection algorithm shave been designed. for categorical data. There are two main problems of outlier detection for categorical data, which are the similarity measure between categorical data objects and the detection efficiency. Outlier detection algorithm for categorical data. Efficient outlier detectioncan help us make good decisions on erroneous data or prevent the negative influence of malicious and faulty behavior. Many data mining techniques try to reduce the influence of outliers or eliminate them entirely. The in foremention manner may result in the loss of important hidden information.[4][5]

Cloud Computing-Based Forensic Analysis for Collaborative Network Security Management System: Zhen Chen*, Fuye Han, Junwei Cao, Xin Jiang, and Shuo Chen.1, February 2013. Internet security problems remain a major challenge with many security concerns such as Internet worms, spam, and phishing attacks. Botnets, well-organized distributed network attacks, consist of a large number of bots that generate huge volumes of spam or launch Distributed Denial of Service (DDoS) attacks on victim hosts. A distributed security overlay network with a centralized security center leverages a peer-to-peer communication protocol used in the UTMs collaborative module. These new security rules are enforced by collaborative UTM and the feedback events of such rules are returned to the security center. Collaborative network security management system can not identify the intrusion.[6] Understanding privacy in data mining requires understanding how privacy can be violated and the possible means for preventing privacy violation. In general, one major factor contributes to privacy violation in data mining: the misuse of data. Users' privacy can be violated in different ways and with different intentions. Although data mining can be extremely valuable in many applications (e.g., business, medical analysis, etc), it can also, in the absence of adequate safeguards, violate informational privacy. Privacy .

III. Flow Chart


•1)Browses files,Encrypts and uploads onto storage node.
•2)Gives privaleges
•Service Provider
•STORAGE NODE
•1) Create End User
•2) Stores Service provides files with secrete key
•1)Logins and download files
•2)Acts as an internal attacker
•Get the Key from key authorities
•End User
•Request for secret key of particular file
•Request for Secrete Key to Key Authorities
•KA1
•KA3



+Service provider
+Encrypt and upload data
+Request and get the key for file
+Storage Area Network
+Finding all end user requested file keys
+Response
+Key Authorities
+Decrypt all related data sets and reply
+Request
+End user

Fig 1. Dataflow diagram

The A typical setting involves two parties: one that seeks information from the other that is either motivated, or compelled, to share the requested information. Consequently, in numerous occasions, there is a tension between information sharing and privacy. On the one hand, sensitive data needs to be kept confidential; on the other hand, data owners may be willing, or forced, to share information.

IV. Problem Detected

The evaluate the accuracy of our solution with several types of datasets under a multitude of data leak scenarios. This module allows the user to register their identity into the system with proper input parameters. The key generation centers play a vital role in it, which generates secret parameters . The key authorities consist of a central authority and multiple local authorities. taking account that there are secure and trusted communication between super authority and each local authority during the initial key and key generation . Each local authority gets different attributes and issues corresponding attribute keys to users. They give differential access rights to individual users based on the users' attributes. The key authorities are provide secured key. That is, they will execute the assigned tasks in the system they would like to information of encrypted contents.

V. Proposed System

The purpose of this proposed work is to provide the functions as a privacy to protect parties from disclosing more than the required minimum of their respective sensitive information. Usage for prompts many problem, The experimental results attest to the practicality of achieved privacy features and show that our approach incurs quite low overhead. For efficient attack detection, big data incorporates attack analytical procedures into the data leak detection processes. There is a note that the does not intend to improve any of the existing data content leakage algorithms indeed, employs a reconfigurable virtual networking approach to detect and counter the attempts to compromise VMs, thus preventing zombie VMs.

The proposed method has several advantages.

1. To avoid the attacker.
2. Secrecy of the data should be maintained.
3. The scheme is robust to withstand brute force attacks.

The privacy in data mining requires to how privacy can be violated and the any means for preventing privacy violation. One main factor contributes to privacy violation in data mining the misuse of data. Users' privacy can be violated in different ways and with different intentions. the not proper adequate safe protects, violate informational privacy. Privacy can be violated if personal data are used for other purposes subsequent to the original transaction between an individual and an organization when the information was collected (Culnan, 1993).

One of the sources of privacy problem is known as data magnets (Rezgui et al., 2003). Data magnets are method and tools used to collect personal data. The method include explicitly collecting information through registration, identifying users through IP addresses, software downloads that require registration, and indirectly getting the information for secondary usage. In many cases, users may or may not be aware that information is being collected or do not know how that information is collected. The collected user data can be used for secondary usage beyond the users' access control and privacy laws. This has led to an uncontrollable privacy violation not because of data mining, but fundamentally because of the misuse of data.

- ***Individual privacy preservation*****:** The primary goal of data privacy is the protection of personally identifiable information. The information is considered personally identifiable if it can be linked, directly or indirectly, to an individual person. when personal data are taken in for to mining, the attribute values linked with individuals are private and must be protected from disclosure. Miners are then able to learn from global models rather than from the characteristics of a particular individual.

- ***Collective privacy preservation*****:** Protecting personal data may not be enough. There is necessary to protect data against sensitive knowledge representing the activities of a group. The protection of sensitive knowledge as collective privacy preservation. The goal here is similar to that one for databases, in which security control mechanisms provide aggregate information about groups and, at the same time, prevent disclosure of confidential information about individuals. It, unlike as is the case for databases, the main objective of collective privacy preservation is to protect sensitive knowledge that can provide competitive advantage in the commercial world.

In the case of collective privacy preservation, organizations have to scope with some interesting problem. The information analysis processes that produce new facts about users' shopping patterns, hobbies, or preferences, these facts could be used in recommender systems to predict or affect their future shopping patterns. This scenario is beneficial to both users and organizations. When the share data in a collaborative system, the aim is not only to protect personally identifiable information but also sensitive data knowledge represented by some strategic patterns. To increase the security level this proposed scheme overcomes the limitation of “Hybrid encryption algorithm proposed . The proposed enhanced scheme has Triple DES, MD5and RSA. Triple DES strengthens the security of Data transmission. The purpose behind for selecting triple DES rather than Double DES is that in double DES algorithm the key used for encryption and decryption is suspected to middle man attack. key distribution problem and in addition to this, MD5 to verify the data of the message. The message digest algorithm in combination of cryptographic algorithm.

VI. Research Method

1) Identity Key Generation The key generation module helps the users to share the information between source and destination. After getting the confirmation response from the receiver side the sender fix the information and encrypt it. At this time a key will be generated and sent to the receiver area. That key is useful for decrypt the data at receiver end. As well as an entity that stores data from senders and provide corresponding access to users. It may be mobile or static. Similar to the previous schemes, and also assume the storage node to be semi trusted that is honest but curious.
2) 3DES Based Encryption In Ciphertext Policy Attribute based Encryption scheme, the encryptor can fix the policy, who can decrypt the encrypted message. The policy can be formed with the help of attributes. In access policy is sent along with the ciphertext. We propose a method in which the access policy need not be sent along with the ciphertext, by which we are able to preserve the privacy of the encryptor. This method encrypted data can be kept confidential even if the storage server is untrusted. Moreover, our methods are secure against collusion attacks. Previous Attribute Based Encryption systems used attributes to describe the encrypted data and built policies into user's keys; while in our system attributes are used to describe a user's value, and a party encrypting data has a policy for that can decrypt.
3) Confidential Data Interchange This is an entity who owns confidential messages or data and wishes to store them into the external data storage node for ease of sharing or for important delivery to users in the networking environments. A sender is responsible for defining (attribute based) access policy and enforcing it on its own data by encrypting the data under the policy before storing it to the storage node. This is a mobile node who wants to access the data stored at the storage node (e.g., a soldier). If a user possesses a set of attributes satisfying the access policy of theencrypted data defined by the sender, and is not revoked in any of the attributes, then he will be able to decrypt the ciphertext and obtain the data.
4) Administrative Access Controller The administrator owns full access rights of this entire site. Once the administrator find out any illegal activity or other misusing happens into the way of transaction between the respective sender and receiver then the admin immediately block the user access rights to transact using this site. The block will be unblocked after getting meaningful reason from the user end.

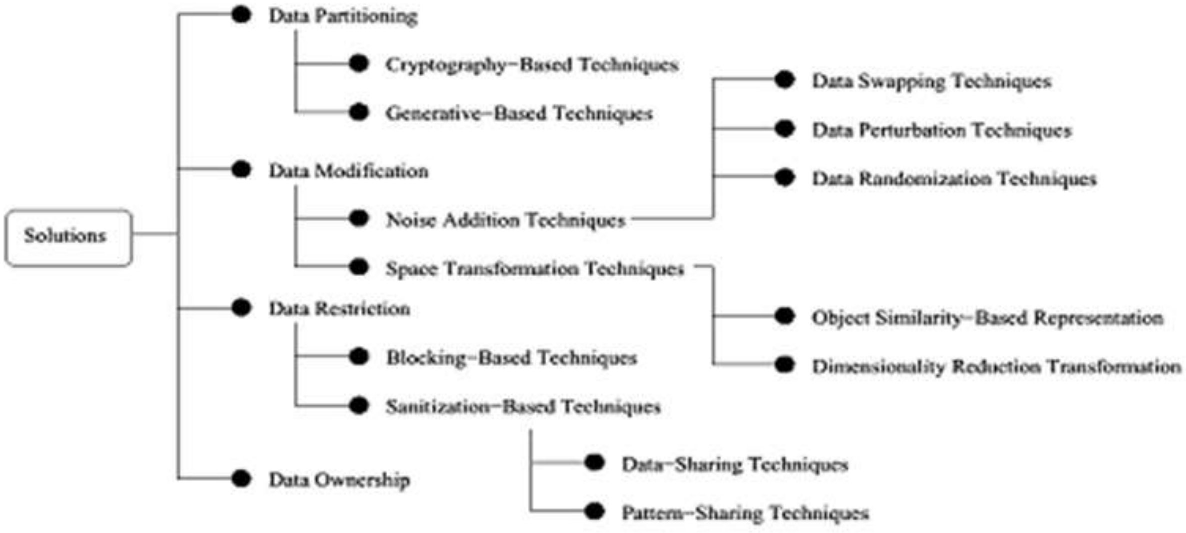
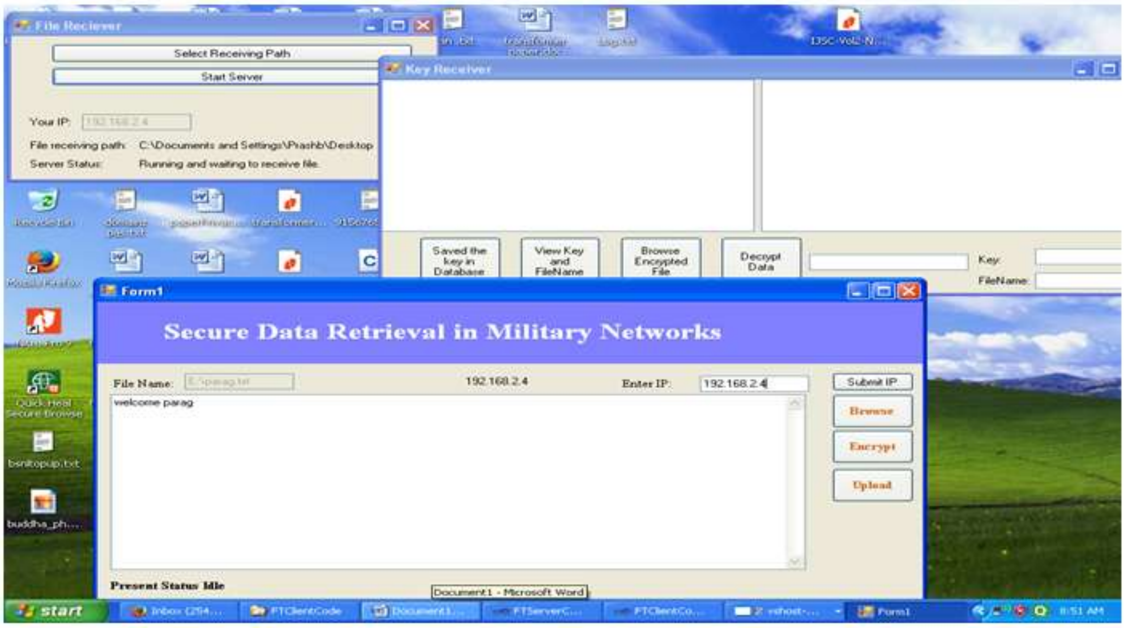

Solutions
Data Partitioning
Cryptography–Based Techniques
Generative–Based Techniques
Data Modification
Noise Addition Techniques
Space Transformation Techniques
Data Restriction
Blocking–Based Techniques
Sanitization–Based Techniques
Data Ownership
Data Swapping Techniques
Data Perturbation Techniques
Data Randomization Techniques
Object Similarity–Based Representation
Dimensionality Reduction Transformation
Data–Sharing Techniques
Pattern–Sharing Techniques

Figure 2 A taxonomy of PPDM techniques

VII. DES With Md5 Algorthim

3DES encrypts a 64-bit block of plaintext to 64-bit block of ciphertext. It usesa 128-bit key. The algorithm consists of eight identical rounds and a "half" roundfinal Transformation. There are 216 possible 16-bit blocks: ,. Each operation with the set of possible 16-bit blocks is an algebraic group. Bitwise XOR is bitwise addition modulo 2,and addition modulo 216 is the usual group operation. Some spin must be put on the elements – the 16-bit blocks – to make sense of multiplication modulo 216 + 1, however. 0 is not an element of the multiplicative group.



Your IP: 192.168.2.4

File receiving path: C:\Documents and Settings\Prashb\Desktop

Server Status: Running and waiting to receive file.

Secure Data Retrieval in Military Networks

File Name: E:\parag.txt 192.168.2.4 Enter IP: 192.168.2.4

welcome parag

Present Status Idle

Figure 3. Secured Data Retrieval System

Confidentiality: In order to protect sensed data and communication exchanges between sensor nodes it is important to guarantee the secrecy of messages. In the sensor network case this is usually achieved by the use of symmetric cryptography as asymmetric or public key cryptography in general is considered too expensive. However, while encryption protects against outside attacks, it does not protect against inside attacks/node compromises, as an attacker can use recovered cryptographic key material to successfully eavesdrop, impersonate or participate in the secret communications of the network. Furthermore, while confidentiality guarantees the security of communications inside the network it does not prevent the misuse of information reaching the base station. Hence, confidentiality must also be coupled with the right control policies so that only authorized users can have access to confidential information.Integrity and **Authentication:** authentication is necessary to enable sensor nodes to detect modified, injected, or replayed packets. While it is clear that safety-critical applications require authentication, it is still wise to use it even for the rest of applications since otherwise the owner of the sensor network may get the wrong picture of the sensed world thus making inappropriate decisions. The authentication does not solve the problem of node takeovers as compromised nodes can still authenticate to the network. Hence authentication mechanisms should be "collective" and aim at securing the entire network.In particular, the following requirements must be supported by the key management scheme, in order to facilitate data aggregation and dissemination process: First we focused on the establishment of trust relationship among wireless sensor nodes, and presented a key management protocol for sensor networks. The protocol includes support for establishing four types of keys per sensor node: individual keys shared with the base station, pairwise keys shared with individual neighboring nodes, cluster keys shared with a set of neighbors, and a group key shared with all the nodes in the network. We showed how the keys can be distributed so that the protocol can support in-network processing and efficient dissemination, while restricting the security impact of a node compromise to the immediate network neighborhood of the compromised node. Applying the protocol makes it really hard for an adversary to disrupt the normal operation of the network.

VIII. Conclusion

The corresponding attribute group keys are updated and delivered to the valid attribute group members securely. In addition, all of the components encrypted with a secret key in the ciphertext are reencrypted by the storage node with a random, and the ciphertext components corresponding to the attributes are also reencrypted with the updated attribute group keys. Even if the user has stored the previous ciphertext exchanged before he obtains the attribute keys and the holding attributes satisfy the access policy, he cannot decrypt the pervious ciphertext..

References

[1]. Hiroki Nishiyama, Desmond Fomo, Zubair Md. Fadlullah,, and NeiKato,Fellow," Traffic Pattern Based Content Leakage Detection for Trusted Content Delivery Networks" IEEE Transaction on Parallel and Distributed System , Volume 25, No 2 Feb 2014

[2]. K. Ramya, D. RamyaDorai, Dr. M. Rajaram "Tracing Illegal Redistributors of Streaming Contents using Traffic Patterns" IJC A 2011

[3]. A. Asano, H. Nishiyama, and N. Kato, "The Effect of Packet Reordering and Encrypted Traffic on Streaming Content Leakage Detection" Proc. Int'l Conf. Computer Comm. Networks (ICCCN '10), pp. 1 6, Aug. 2010.

[4]. Y. Chu, S.G. Rao, S. Seshan, and H. Zhang, "Enabling Conferencing Applications on the Internet Using an Overlay Multicast Architecture," Proc.ACM SIGCOMM, pp. 55 67,Aug. 2010

[5]. O. Adeyinka, "Analysis of IPSec VPNs Performance in a Multimedia Environment," Proc. Fourth Int'l Conf. Intell igent Environments, pp. 25 - 30, 2008

[6]. M. Dobashi, H. Nakayama, N. Kato, Y. Nemoto, and A. Jamalipour, "Traitor Tracing Technology of Streaming Contents Delivery Using Traffic Pattern in Wired/Wireless Environments," Proc. IEEE Global Telecomm. Conf., pp. 1 5, Nov./Dec. 2006.

[7]. S. Amarasing and M. Lertwatechakul, "The Study of Streaming Traffic Behavior," KKU Eng. J., vol. 33, no. 5, pp. 541 553, Sept./Oct. 2006.

[8]. R.S. Naini and Y. Wang, "Sequential Traitor Tracing," IEEE Trans. Information Theory, vol. 49, no. 5, pp. 1319 1326, May 2003.

[9]. D. Geiger, A. Gupta, L.A. Costa, and J. Vlontzos, "Dynamic Programming for Detecting, Tracking, and Matching Deformable C ontours," Proc. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 17, no. 3, pp. 294 302, M ar. 1995

Examples follow:

Journal Papers:

[10]. M Ozaki, Y. Adachi, Y. Iwahori, and N. Ishii, Application of fuzzy theory to writer recognition of Chinese characters, *International Journal of Modelling and Simulation, 18(2),* 1998, 112-116. (8)
Note that the journal title, volume number and issue number are set in italics.

Books:

[11]. R.E. Moore, *Interval analysis* (Englewood Cliffs, NJ: Prentice-Hall, 1966). (8)
Note that the title of the book is in lower case letters and italicized. There is no comma following the title. Place of publication and publisher are given.

Chapters in Books:

[12]. P.O. Bishop, Neurophysiology of binocular vision, in J.Houseman (Ed.), *Handbook of physiology,* 4 (New York: Springer-Verlag, 1970) 342-366. (8)

Note that the place of publication, publisher, and year of publication are enclosed in brackets. Editor of book is listed before book title.

Theses:

[13]. D.S. Chan, *Theory and implementation of multidimensional discrete systems for signal processing*, doctoral diss., Massachusetts Institute of Technology, Cambridge, MA, 1978. (8)
Note that thesis title is set in italics and the university that granted the degree is listed along with location information

Proceedings Papers:

[14]. W.J. Book, Modelling design and control of flexible manipulator arms: A tutorial review, *Proc. 29th IEEE Conf. on Decision and Control*, San Francisco, CA, 1990, 500-506 (8)