INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 2 (2018) PAGES 3014 - 3020
RECEIVED : 28.02.2018 PUBLISHED : 29.03.2018
March 29, 2018

DATA LEAKAGE DETECTION

1.Mrs. R.Angeline 2.Paruchuri Yaswanth 3.L. Ali Khan 4.Sk. Ashar 5.P.Santhosh

1 Assistant Professor, Department of Computer Science Engineering, SRM IST Ramapuram, Chennai - 600089.

2 Student, Department of Computer Science Engineering, SRM IST Ramapuram, Chennai-600089.

ABSTRACT

A data distributor has given sensitive data to a set of supposedly trusted agents (third parties). Some of the data is leaked and found in an unauthorized place (e.g., on the web or somebody's laptop). The distributor must assess the likelihood that the leaked data came from one or more agents, as opposed to having been independently gathered by other means. We propose data allocation strategies (across the agents) that improve the probability of identifying leakages. These methods do not rely on alterations of the released data (e.g., watermarks). In some cases we can also inject "realistic but fake" data records to further improve our chances of detecting leakage and identifying the guilty party.

Network security states that the data or information or resource that is available on the network can only be accessed by authorized user and it also prevents the data or information or resource from unauthorized user from malicious practice. Network administrator is the person who provides authorization for users to access their data. There are two types of network available. In Private network, Security is provided within an organization or company. In Public network, Security is providing globally by the use of user name and password. Only the person who has the authorization can use it and others are restricted. Network security usually starts with three factors of authentication facilities are one-factor authentication, two-facto authentication, and three-factor authentication. Firewall acts as a protection wall between the user and the network and allows only authorized user to access their information or data or resources over the network The collected data provide the information about the various environmental factors. Monitoring the environmental factors is not the complete solution to increase the yield of crops. There are number of other factors that decrease the productivity. Hence automation must be implemented in irrigating fields to overcome these problems. So, to provide solution to all such problems, it is necessary to develop an integrated system which will take care of watering the crops. But complete automation in irrigation is not achieved due to various issues. Though it is implemented in the research level it is not given to the farmers as a product to get benefitted from the resources. Hence this paper deals about Automatic Irrigation System using Iot.

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 2 (2018) PAGES 3014 - 3020
RECEIVED : 28.02.2018 PUBLISHED : 29.03.2018
March 29, 2018

1. INTRODUCTION

Data leakage is the unauthorized transmission of data or information from within an organization to an external destination or recipient . Data leakage is defined as the accidental or intentional distribution of private or sensitive data to an unauthorized entity. Sensitive data of companies and organization includes intellectual property, financial information, patient information, personal credit card data and other information depending upon the business and the industry. Furthermore, in many cases, sensitive data shred among various stakeholders such as employees working from outside the organizational premises, business partners and customers. This increases the risk of confidential information falling into unauthorized hands.

Furthermore, in many cases, sensitive data is shared among various stakeholders such as employees working from outside the organizational premises (e.g., on laptops), business partners and customers. This increases the risk of confidential information falling into unauthorized hands .In the course of doing business, sometimes data must be handed over to supposedly trusted third parties for some enhancement or operations

2. LITERATURE SURVEY

Network security states that the data or information or resource that is available on the network can only be accessed by authorized user and it also prevents the data or information or resource from unauthorized user from malicious practice. Network administrator is the person who provides authorization for users to access their data. There are two types of network available. In Private network, Security is provided within an organization or company. In Public network, Security is providing globally by the use of user name and password. Only the person who has the authorization can use it and others are restricted. Network security usually starts with three factors of authentication facilities are one-factor authentication, two-factor

authentication, and three-factor authentication. Firewall acts as a protection wall between the user and the network and allows only authorized user to access their information or data or resources over the network The collected data provide the information about the various environmental factors. Monitoring the environmental factors is not the complete solution to increase the yield of crops. There are number of other factors that decrease the productivity. Hence automation must be implemented in irrigating fields to overcome these problems. So, to provide solution to all such problems, it is necessary to develop an integrated system which will take care of watering the crops. But complete automation in irrigation is not achieved due to various issues. Though it is implemented in the research level it is not given to the farmers as a product to get benefitted from the resources. Hence this paper deals about Automatic Irrigation System using Iot.

Data leakage is the unauthorized transmission of data or information from within an organization to an external destination or recipient . Data leakage is defined as the accidental or intentional distribution of

private or sensitive data to an unauthorized entity. Sensitive data of companies and organization includes intellectual property, financial information, patient information, personal credit card data and other information depending upon the business and the industry. Furthermore, in many cases, sensitive data shred among various stakeholders such as employees working from outside the organizational premises, business partners and customers. This increases the risk of confidential information falling into unauthorized hands.

Furthermore, in many cases, sensitive data is shared among various stakeholders such as employees working from outside the organizational premises (e.g., on laptops), business partners and customers. This increases the risk of confidential information falling into unauthorized hands .In the course of doing accurate data for the patients . Traditionally, leakage detection is handled by the watermarking. For example a unique code is embedded in each distributed copy. If that copy is later found in the hands of an unauthorized party, the leaker can be identified. Watermarks can be very useful in some cases but again, involve some modification of the original data. Furthermore, watermarks can sometimes be destroyed if the data recipient is malicious [research paper].

In short, watermarking is suitable for all the application because it's lost the original data. There are some disadvantages of it. That is It involves some modification of data that is making the data less sensitive by altering attributes of the data. The second problem is that these watermarks can be sometimes destroyed if the recipient is malicious .In this paper, we develop an algorithm of data allocation strategies for finding the guilty agents that improves the chances of identifying a leaker. We also consider the option of adding fake objects to the distributed set. Such object do not corresponds to real entities but appear realistic to the agents. Means that fake objects act as a type of watermarks for the entire set, without modifying any original data. If it turns out that an agent was given one or more fake objects that were leaked, then the distributor can be more confident that agent was guilty.

3. EXISTING SYSTEM

Traditionally, leakage detection is handled by watermarking, e.g., a unique code is embedded in each distributed copy. If that copy is later discovered in the hands of an unauthorized party, the leaker can be identified. Watermarks can be very useful in some cases, but again, involve some modification of the original data. Furthermore, watermarks can sometimes be destroyed if the data recipient is malicious. E.g. A hospital may give patient records to researchers who will devise new treatments. Similarly, a company may have partnerships with other companies that require sharing customer data. Another enterprise may outsource its data processing, so data must be given to various other companies. We call the owner of the data the distributor and the supposedly trusted third parties the agents. We consider applications where the original sensitive data cannot be perturbed. Perturbation is a very useful technique where the data is modified and made "less sensitive" before being handed to agents .However, in some cases it is important not to alter the original distributor's data

.Traditionally, leakage detection is handled by watermarking, e.g., a unique code is embedded in each distributed copy .If that copy is later discovered in the hands of an unauthorized party, the leaker can be identified .Watermarks can be very useful in some cases, but again, involve some modification of the original data

Furthermore, watermarks can sometimes be destroyed if the data recipient is malicious. means that some third party called the target has been caught in possession of S. For example, [9] this target may be displaying S on its web site, or perhaps as part of a legal discovery process, the target turned over S to the distributor. Since the agents (A1, A2,…An) have some of the data, it is reasonable to suspect them leaking the data. However, the agents can argue that they are innocent, and that the S data was obtained by the target through other means.

Example: Say T contains customer records for a given company A. Company A hires a marketing agency U1 to do an on-line survey of customers. Since any customers will do for the survey, U1 requests a sample of 1000 customer records. At the same time, company A subcontracts with agent U2 to handle billing for all California customers. Thus, U2 receives all T records that satisfy the condition "state is California."
Guilty Agents

Suppose that after giving objects to agents, the distributor discovers that a set S _ T has leaked. This means that some third party called the target, has been caught in possession of S. For example, this target may be displaying S on its web site, or perhaps as part of a legal discovery process, the target turned over S to the distributor. Since the agents U1; : : : ;Un have some of

4. ISSUES IN EXISTING SYSTEM

Entities and Agents Let the distributor database owns a set S= {t1, t2,…., tm} which consists of data objects. Let the no of agents be A1, A2, ..., An [6][10]. The distributor distributes a set of records S to any agents based on their request such as sample or explicit request. Sample request Ri= SAMPLE (T, mi): Any subset of mi records from T can be given to Ui [1].Explicit request Ri= EXPLICIT (T;condi):Agent Ui receives all T objects that satisfy Condition .The objects in T could be of an type and size, e.g. they could be tuples in a relation, or relations in a database. After giving object to agents, the distributor discovers that a set S of T has leaked. This
the data, it is reasonable to suspect them leaking the data. However, the agents can argue that they are innocent, and that the S data was obtained by the target through other means. For example, say one of the objects in S represents a customer

X. Perhaps X is also a customer of some other company, and that company provided the data to the target. Or perhaps X can be reconstructed from various publicly available sources on the web.

Our goal is to estimate the likelihood that the leaked data came from the agents as opposed to other sources. Intuitively, the more data in S, the harder it is for the agents to argue they did not leak anything.

Similarly, the "rarer" the objects, the harder it is to argue that the target obtained them through other means. Not only do we want to estimate the likelihood the agents leaked data, but we would also like to find out if one of them in particular was more likely to be the leaker. For instance, if one of the S objects was only given to agent U1, while the other objects were given to all agents, we may suspect U1 more. The model we present next captures this intuition. We say an agent Ui is guilty if it contributes one or more objects to the target. We denote the event that agent Ui is guilty for a given leaked set S by GijS. Our next step is to
estimate PrfGijSg, i.e., the probability that agent Ui is guilty given evidence S.

5. PROPOSED SYSTEM

A data distributor has given sensitive data to a set of supposedly trusted agents (third parties). Some of the data is leaked and found in an unauthorized place (e.g., on the web or somebody's laptop). The distributor must assess the likelihood that the leaked data came from one or more agents, as opposed to having been independently gathered by other means. We propose data allocation strategies (across the agents) that improve the probability of identifying leakages.

These methods do not rely on alterations of the released data (e.g., *watermarks*). In some cases we can also inject "realistic but fake" data records to further improve our chances of detecting leakage and identifying the guilty party.

6. CONCLUSION

In a perfect world there would be no need to hand over sensitive data to agents that may unknowingly or maliciously leak it. And even if we had to hand over sensitive data, in a perfect world we could watermark each object so that we could trace its origins with absolute certainty. However, in many cases we must indeed work with agents that may not be 100% trusted, and we may not be certain if a leaked object came from an agent or from some other source. In spite of these difficulties, we have shown it is possible to assess The likelihood that an agent is responsible for a leak, based on the overlap of its data with the leaked data and the data of other agents, and based on the probability that objects can be "guessed" by other means. Our model is relatively simple, but we believe it captures the essential trade-offs.

7. REFERENCES

[1] K. Lakshmisudha, Swathi Hegde, Neha Kale, Shruti Iyer, "Smart Precision Based Agriculture Using Sensors", International Journal of Computer Applications (0975- 8887), Volume146-No.11, July 2011

[2] Nikesh Gondchawar, Dr. R.S. Kawitkar, "IoT Based Smart Agriculture", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), Vol.5, Issue 6, June 2016.

[3] M.K. Gayatri, J. Jayasakthi, Dr.G.S. Anandhamala, "Providing Smart Agriculture Solutions to Farmers for Better YieldingUsingIoT",IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development (TIAR 2015).

[4] Chetan Dwarkani M, Ganesh Ram R, Jagannathan S, R. Priyatharshini, "Smart Farming System UsingS ensorsfor Agricultural Task Automation", IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development (TIAR 2015).

[5] S. R. Nandurkar, V. R. Thool, R. C. Thool, "Design and Development of Precision Agriculture System Using Wireless Sensor Network", IEEE International Conference on Automation, Control, Energy and Systems (ACES), 2014.

[6] JoaquínGutiérrez,Juan FranciscoVilla-Medina, Alejandra Nieto-Garibay, and Miguel Ángel Porta- Gándara, "Automated Irrigation System Using a Wireless Sensor Network and GPRS Module", IEEE Transactions on Instrumentation and Measurements, 0018-9456,2013

[7] Dr. V. Vidya Devi, G. Meena Kumari, "Real- Time Automation andMonitoringSystemfor ModernizedAgriculture", International Journal of Review and Research in Applied Sciences and Engineering (IJRRASE) Vol3 No.1. PP 7-12, 2013.

[8] MeonghunLee, Jeonghwan Hwang, Hyun Yoe, "Agricultural Protection System Based IoT", IEEE 16th International Conference on Computational Science and Engineering, 2013.

[9] Monika Jhuria, Ashwani Kumar, RushikeshBorse, "Image Processing for Smart Farming: Detection of Disease and Fruit Grading", IEEE Second International Conference on Image Information Processing (ICIIP), 2013.

[10]Orazio Mirabella and Michele Brischetto,"AHybrid Wired/Wireless Networking Infrastructure for Greenhouse Management", IEEE Transactions Instrumentation and Measurement, vol. 60, no. 2, pp 398407, 2011.

[11]C. Liu, W. Ren, B. Zhang, and C. Lv, "The application of soil temperature measurement by lm35 temperature sensors," International Conference on Electronic and Mechanical Engineering and Information Technology, vol. 88, no. 1, pp. 1825–1828, 2011

[12] Q. Wang, A. Terzis and A. Szalay, "A Novel Soil Measuring Wireless Sensor Network", IEEE Transactions on Instrumentation and Measurement, pp. 412–415, 2010

[13] Ji-woong Lee, Changsun Shin, Hyun Yoe," An Implementation of Paprika Greenhouse System Using Wireless Sensor Networks", International Journal of Smart Home Vol.4, No.3, July 2010.