Paper 1783-2014

Revealing Unwarranted Access to Sensitive Data: A Scenario-based Approach

Heidi Thorstensen (uxthei@ous-hf.no)

ICT Security Manager, Oslo University hospital, Norway

Torulf Mollestad Ph.D. (Torulf.mollestad@sas.com),

Principal Advisor, Enterprise Architecture CoE, SAS® Institute Nordic

Abstract

The project focuses on using analytics to reveal unwarranted use of access to medical records, i.e. employees in health organizations that access information about neighbours, friends, celebrities, etc., without a sound reason to do so. The method is based on the natural assumption that the vast majority of lookups are legitimate – lookups that differ from a statistically defined normal behavior will be subject to manual investigation. The work was carried out in collaboration between SAS® Institute Norway and the largest Norwegian hospital, Oslo University Hospital (OUS) and was aimed at establishing whether the method is suitable for unveiling unwarranted lookups in medical records.

A number of so called scenarios are used to indicate adverse behaviour, each responsible for looking at one particular aspect of journal access data. For instance, one scenario determines the timeliness of a lookup relative to the patient's admission history; another judges whether the medical competency of the employee is relevant to the situation of the patient at the time of the lookup. We have so far designed and developed a library of around 20 scenarios that together are used in weighted combination to render a final judgment of the appropriateness of the lookup. The approach has been proven highly successful, and a further development of these ideas is currently being done, the aim of which is to establish a joint Norwegian solution to the problem of unwarranted access. Furthermore, we believe that the approach and the framework may be utilised in many other industries where sensitive data is being processed, such as financial, police, tax and social services. In this paper, the method is outlined, as well as results of its application on data from OUS.

Introduction

Are your personal data safe? Who has access to your information, and what interests are served in looking at it?

We are in the process of building up a large collective memory. Each person in developed countries today is enrolled in a hundred different registers and databases, information about our medical history from birth to death, ditto for our economic history and financial problems. Conflicts with police and legal system are duly recorded, and the same applies to information about our professional career and any periods of unemployment.

Illegitimate access to information can provide fertile ground for fraud and crime, not to mention loss of reputation for the company or for individuals. Employees of different agencies have access to sensitive information about celebrities, politicians, relatives, neighbors and others; control over how that information is used is at best lacking. It is assumed that the employees have strong integrity and do not violate internal policies for access and sharing of sensitive

information – for purposes of control, however, this is of course not good enough. Whereas businesses traditionally have been quick to hedge against external threats and data breaches, they are often less vigilant toward potential internal threats. There is an increasing tendency for criminals to get into contact with company employees and to get them, in exchange of compensation or under threat, to disclose confidential information. An example of this is a recent record fraud in a large Nordic bank, where a trusted employee was found to be very centrally placed in the criminal conspiracy. Due to poor internal routines, the bank fell victim to an all-time record scam.

We believe that the main remedy against unwarranted data access is found in the form of automatic monitoring algorithms. In this paper we describe a solution that aims to reveal unwarranted lookups on patient data, developed under a cooperation between SAS Institute and several hospitals in Norway, and especially Oslo University Hospital (OUS).

The Hospital and the situation

All Norwegian hospitals are obliged to maintain control of the treatment of patient information and how that information is utilised. It is required by law that each lookup in medical records be logged and also that each hospital reviews its logs to verify that the lookups were performed on a legitimate basis. However, most hospitals do not have proper routines or tools in place to do a meaningful investigation. When doing inspections, they typically prioritise cases where patients themselves suspect malpractice and make a complaint, or they monitor activity against celebrities of different sorts. As a last resort, they may have to rely on pure random selection and manual inspection. It is evident that the latter is bound to fail, considering that hospitals typically have tens of millions of lookups each year. The first two situations raise another pertinent question, whether celebrities or people of a nervous disposition should be gratified with a higher level of safety for their private data than is the case for the rest of us.

Oslo University Hospital is a highly specialized hospital in charge of extensive regional and local hospital assignments and the provision of high quality services to the citizens of Oslo. The hospital also has a nationwide responsibility for a number of national and multi-regional assignments and has several national centers of competence. The hospital is the largest in Scandinavia, carrying out more than 1.2 million patient treatments each year. OUS is an emergency hospital for East and Southern Norway and also handles emergency assignments on a national level.

The amount of sensitive information is vast. The logs contain more than 100 million lookups per year. In addition, controlling access by setting strict restrictions in the system will easily conflict with patient security. An emergency hospital must allow medical records to be easy and quick to access at any time for a large number of personnel. Modern health care is often performed across traditional professional boundaries, making it impossible to predetermine which doctors need a wider access than do other colleagues.

The result is that all doctors, with very few exceptions, may access any medical record in the hospital. In many cases, this also implies nurses and administrative personnel. This situation reflects the many different legal grounds for accessing medical records, in great extent leaving it to the personnel to consider whether a lookup is legal or not. In other words; to avoid conflict

with patient security, the technical ability to access a medical record is often wider than the actual daily professional need would warrant.

Likewise, it is considered impossible to analyze logs based on predetermined rules. The complexity of health care in a large organization simply does not permit such an approach, and that is why OUS chose to investigate the method described in this paper.

The method and solution

The method works by gathering circumstantial evidence around individual lookups, the people that make them and the people that are looked at. Many aspects may help determine whether a lookup is likely to be legitimate or not. For instance, holding the patient's medical history, referrals and admissions up against the employee's skills and job description may reveal patterns of obvious inconsistency. The time of the lookup vs. the situation of the patient may also be telling, for instance a lookup three years after the patient was admitted from the hospital may be an indication of possible misconduct, whereas a lookup done at a time when the patient was undergoing surgery may seem less severe – *considering that piece of information only.*

In the solution, each aspect of a given lookup (or properties of the people involved) is covered by a so called scenario. In the course of this project, we have developed a library of around 20 different scenarios, each of which is responsible for but one little piece of the total picture. Some scenarios are based on the properties of the lookup itself. For instance, the timeliness of the lookup is covered by the *time scenario (Fig. 1).* Assume that “Chris” is admitted to the hospital; a so-called *admission period* starts at the time that the hospital receives information about Chris’ condition, and ends when they are no longer responsible for the treatment of this diagnosis. During the period, there are certain points where Chris has a bit closer contact with the hospital, for instance a call made to him to check up, a visit to the doctor or even undergoing surgery. Such time points or intervals are all found within an admission period, and are called *contacts*.

Now, “Bob” and “Rita” are both employees at the hospital, and when overseeing the lookup records, it is found that Bob accessed Chris’ data a long time after admission (i.e. well outside an admission period). Rita’s lookup on Chris, on the other hand, is in perfect sync with a time that Chris was undergoing surgery (i.e. within an admission period/episode). It would seem reasonable to say that, considering only the time aspect, Bob’s lookup may seem less relevant and therefore more suspicious than Rita’s. When run through the scenario, then, the prior lookup would receive a higher score than the latter. If the lookup is made closely before or after an admission period or episode, the score returned would lie in the middle ground, the level decided by how closely it precedes or follows the actual patient activities.

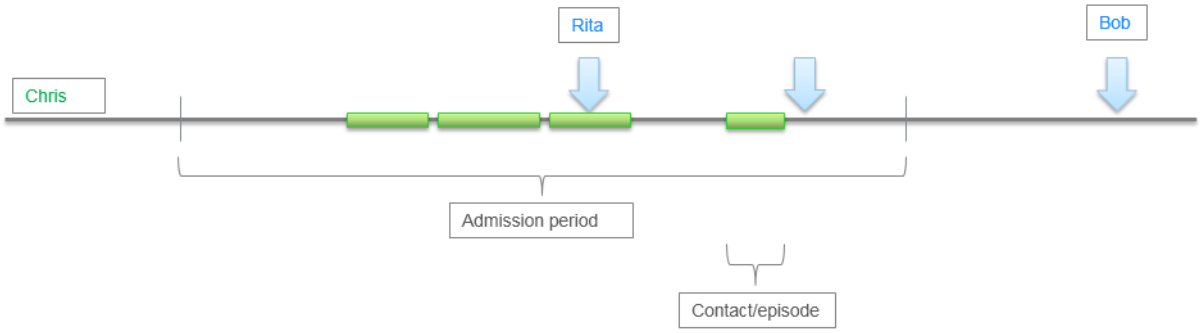

Figure 1: Time of the lookup vs. patient situation

Other scenarios focus on characteristics of the employee and of the patient, or on discrepancies between them. For instance, "*Why did the specialist in geriatrics look at this child patient?*" Consider the situation shown in *Fig. 2* – here we take another perspective to Bob's lookup. Assume that Bob works in department A in the hospital. Typically, people affiliated with that department look at patients admitted to the same department (which is be expected) and we also see that, statistically, the same employees often look at patients that are admitted to Department B and C (this is shown by the shaded big arrows between the departments). Bob follows this pattern too, but he also has looks at people that are admitted to departments that are seldom or never looked at by dept. A – more specifically, Bob looks at Chris who is admitted to a department that is seldom or never looked at by Bob's colleagues. In this light, the activity against Chris' data would seem more unusual and potentially suspicious than would be activities against patients that are admitted to departments that are more consistent with Bob's affiliation with department A.

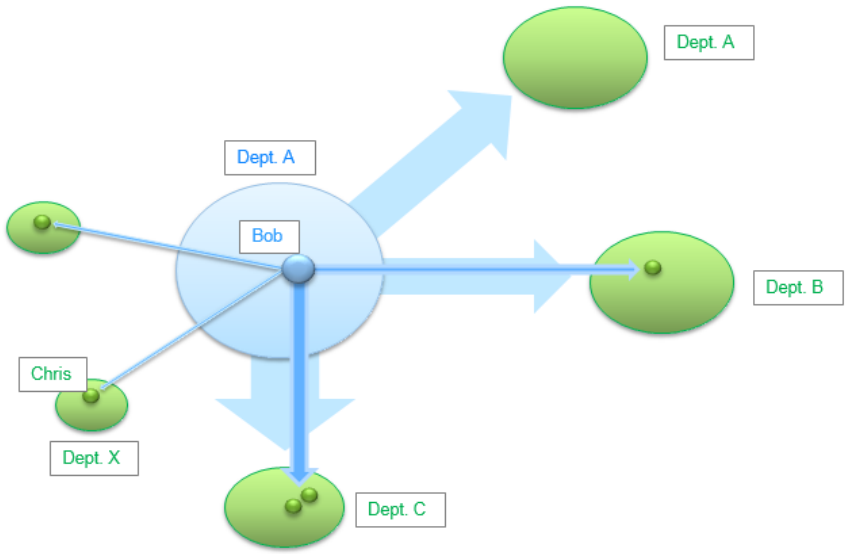

Figure 2: Lookups of the individual vs. the department

At this point, we have considered Bob's lookup at Chris from two angles, namely considering first the time aspect and then the organisational/departmental perspective. Each on its own, either of these two dimensions are in themselves not enough to raise suspicion; the gathering of these circumstantial evidence, however, gives us a stronger argument. In the case above, the lookup would receive high scores on both scenarios and it is therefore bumped up the priority list.

For each lookup, each scenario returns a score between 0 and 1000 that signals the alert level of that lookup. Most lookups, in turn, may generate high scores on a fraction of the scenarios, but that is to be expected and should not be enough to warrant closer inspection. In this light, note that each scenario is in itself not powerful enough to say conclusively that a given lookup represents a breach. However, when all the circumstantial evidence is combined (i.e. scores from up to 20 different scenarios collected), a total score is computed which weights the scenarios. Lookups that are in severe breach of some scenarios will be sent upward the list, even when they may have a very low score on others. For instance, considering ten scenarios, a lookup that scores maximum (1000) on three and minimum (0) on seven of them, will rate higher on the list than a lookup that scores 300 on all, even though the average score is the same for both lookups.

The total picture gives an indication of whether a lookup is likely to be legitimate. Still, a number of scenarios returning high scores is in itself not proof of any malpractice, rather indicative that the lookup exhibits properties that warrant a closer, manual follow up and investigation. This is an important principle of the solution; we have developed tools to make a prioritised list based upon the scenarios, and also methods and reports for gathering evidence to make a final, more informed decision. For OUS it is important to emphasise that it is always the clinical managers, supported by the HR department, that can draw the final conclusion as to whether a lookup is based on legal, professional grounds or not.

Results

One half year of data was used in the project, which comprised several million lookups. From the 150 lookups that were found at the top of the list after the solution had been run (with a chosen set of eight scenarios used), 33 cases were found to be interesting enough to be presented to different clinical managers. Several lookups where determined to be illegal by the managers. Considering that the aim of the project was to prove the validity of the method, this result was highly satisfactory to all parties. Even though the result was regarded a success, there are a number of improvements that can be made, in terms of sharpening the individual scenarios, expansion of the scenario library, improvements to the overall processing of the data as well as development of tools to aid the final investigation phase.

One of the recommendations from the project was to establish a main project with the goal of creating a national library of scenarios, and to further develop and test the method at OUS. The project has been initiated and is led by National ICT, an institution whose responsibility it is to coordinate ICT-related initiatives in the Norwegian specialised health care services and also to be a central agent in realizing national efforts and strategies within the ICT area. The project started January this year and will present its results and recommendations by March 2015. OUS

has high expectations regarding the method and is likely to implement it in the coming years, pending recommendation from the National ICT-project.

Summary and conclusion

The method has proven to be highly suitable for detecting unwarranted use of access to medical records. The strength of the method is that it does not use complex rules, but bases its conclusions on analyzing deviations from normal activity, seen from different angles. Probably no method will ever be “bullet proof”, but this approach is likely to be the most successful within health care institutions. Moreover, we believe that other agents that maintain sensitive data may also benefit from exploring the capabilities of this method, addressing essentially the same issues that exist within fields such as financial, police, taxation or social services industries.