Computer Science Tripos part 2 – Security

Lecture 2 – Other Security Policy Models

Ross J. Anderson

1 The Clark-Wilson policy model

Moving now to the civilian side of computer security, the Clark-Wilson policy model seeks to formalise the principles of accounting security that have accumulated over centuries of experience in commercial bookkeeping (in fact, David Wilson worked for a firm of accountants). As we discussed in 1b, the main security principles in commerce and finance are separation of duties and audit; these are implemented by means of double entry bookkeeping in which each transaction causes equal and opposite entries to be made in two ledgers kept by separate people.

For example, if one of my suppliers pays one of my invoices, then I debit my sales ledger by the amount of his cheque (I no longer have his debt as a current asset) and credit my cash account by the same sum of money. This is an example of a well-formed transaction, or as the Clark-Wilson model has it, a transformation procedure.

In a typical banking system, the software many consist of a number of modules which post transactions to the relevant combinations of ledgers. In order to prevent programmers writing code that amends their own account balances, all code will be subjected to scrutiny by an internal auditor, and to testing by a separate test department. Once it has been approved, it will be run on a production machine that does not have a development environment, but only approved object code and data. In case a clever programmer manages to sneak code past the inspection and test processes, the net effect of all the components of a properly formed transaction must be zero, and balancing routines check this; if they fail, alarms will go off. (Transactions that come in during the day, whether from tellers, ATMs or cheque sorters, are posted to temporary journal files; account balances are only updated during an overnight batch run, which has numerous batch totals and cross checks.)

Thus to add £1000 to his own account, the programmer would have to simultaneously subtract £1000 from other accounts; so he is hopefully in no more privileged a position than a teller who could pass credits and debits to a pair of accounts manually. In such cases, one expects that the losers would complain.

The Clark-Wilson model sets out to formalise and generalise this, and can be summarised as follows. There are special procedures whereby data can be input – turned from an unconstrained data item, or UDI, into a constrained data item, or CDI – and integrity verification procedures to check the validity of any CDI (i.e., the books balance). There are also transformation procedures (TPs) which may be thought of as transactions which preserve balance; in the general formulation, they maintain the integrity of CDIs. They also write enough information to an append-only CDI (the audit trail) for transactions to be reconstructed. Access control is by means of triples (subject, TP, CDI), which are so structured that a dual control policy is enforced. We use the formulation in Amoroso:

1. The system will have an IVP for validating the integrity of any CDI
2. Application of a TP to any CDI must maintain its integrity
3. A CDI can only be changed by a TP
4. Subjects can only initiate certain TPs on certain TPIs
5. CW-triples must enforce an appropriate separation of duty policy on subjects
6. Certain special TPs on UDIs can produce CDIs as output
7. Each application of a TP must cause enough information to reconstruct it to be written to a special append-only CDI
8. The system must authenticate subjects attempting to initiate a TP
9. The system must let only special subjects (i.e., security officers) make changes to authorisation-related lists

The Clark-Wilson paper is appended. Some authors have objected that a suitable TCB to support it doesn't exist yet; others have combined it with Biba, or tried to find a smaller model with similar properties. Yet others have raised theoretical objections to the decidability etc of the integrity validation procedures used to verify CDIs. However, it remains an influential model, as it describes the security policy actually enforced – however haphazardly – in many banks and other large commercial organisations.

2 The Chinese Wall model

Bookkeeping is not the only commercial computer security concern. Sometimes there is also a need to manage conflicts of interest. For example, an advertising agency might have more than one bank, or more than one airline, on its books, and it will not in general be permitted for any of the people who work on one of these accounts to have access to any information on another account. Similar situations arise with accountants, merchant banks and other sectors where the number of ‘blue-chip’ service providers is less than the number of customers willing to engage them.

Such companies often claim to maintain a ‘Chinese Wall’ between those parts of their business which would otherwise lead to a conflict of interest, and a ‘Chinese Wall’ security policy is one which provides such compartments. Each user may have access rights to at most one of a number of resources in a given category.

Clearly, like Clark-Wilson (but unlike the simplest versions of Bell-LaPadula) this involves maintaining state, namely of each user’s association with resources. Where this state is to be managed may require a modicum of care. In some environments, the company might be trusted to enforce the policy; in others, an authenticated CV might suffice; in yet others, care might be needed lest one client discover whether a competitor were about to launch a campaign by measuring the number of staff available to him at any one time.

3 Privacy

There are many applications in which we deal with personal information on data subjects. The topical example is the electronic medical record, but the information kept on us by banks, insurance companies, government agencies and others also falls into this category. The use of such information is governed by data protection legislation, which we shall discuss later; here, we are interested in what technical measures may be taken to prevent leakage.

In late 1995, we developed a security policy model for clinical information systems. It was driven by the experience that the large majority of attacks on medical privacy come through insiders. Put simply, we can live with a situation in which a doctor’s receptionist has access to 2,000 patients’ records: there will be abuse from time to time, but at a tolerably low level. However, if the receptionists of Britain’s 32,000 GPs all had access to the records of all 56,000,000 patients in the country, then abuse would be certain and systematic. The likelihood that a resource will be abused depends on its value and the number of people who have access to it; aggregating personal information into large databases increases both these risk factors at the same time.

An interesting feature of medical confidentiality is the primacy of patient consent. This is stated by the doctors’ regulatory body, the General Medical Council, as:

> Patients have a right to expect that you will not pass on any personal information which you learn in the course of your professional duties, unless they agree.

Thus the goals of our security policy were to enforce the principle of patient consent, and to control aggregation. The policy consists of nine principles, which are:

Principle 1: Access control Each identifiable clinical record shall be marked with an access control list naming the people or groups of people who may read it and append data to it. The system shall prevent anyone not on the access control list from accessing the record in any way

Principle 2: Record opening A clinician may open a record with herself and the patient on the access control list. Where a patient has been referred, she may open a record with herself, the patient and the referring clinician(s) on the access control list

Principle 3: Control One of the clinicians on the access control list must be marked as being responsible. Only she may alter the access control list, and she may only add other health care professionals to it

Principle 4: Consent and notification The responsible clinician must notify the patient of the names on his record's access control list when it is opened, of all subsequent additions, and whenever responsibility is transferred. His consent must also be obtained, except in emergency or in the case of statutory exemptions

Principle 5: Persistence No-one shall have the ability to delete clinical information until the appropriate time period has expired

Principle 6: Attribution All accesses to clinical records shall be marked on the record with the subject's name, as well as the date and time. An audit trail must also be kept of all deletions

Principle 7: Information flow Information derived from record A may be appended to record B if and only if B's access control list is contained in A's

Principle 8: Aggregation control There shall be effective measures to prevent the aggregation of personal health information. In particular, patients must receive special notification if any person whom it is proposed to add to their access control list already has access to personal health information on a large number of people

Principle 9: Trusted Computing Base Computer systems that handle personal health information shall have a subsystem that enforces the above principles in an effective way. Its effectiveness shall be subject to evaluation by independent experts.

The point is that a clinician who keeps personal health information on a system that enforces these principles, or sends it between such systems, may have a reasonable expectation that the record will not end up being leaked.

The full policy can be got from the BMA or from my web page; since it was published in January 96 there have been successful pilots both in general practice and in hospitals, with a compliant system now running in Hastings, Aintree and Exeter. This system taught us firstly, that we might have expressed the policy more concisely in terms of capabilities (a typical implementation rule is that 'a ward nurse can see the records of all patients who have within the previous 90 days been on her ward') and secondly the difficulty of constructing a small TCB (we have to rely on the patient administrative system to tell us which patients, and which nurses, are on which ward). Nonetheless it appears to have made a useful contribution.

Medical privacy is not the only application in which these principles might be used. There are similar concerns with intelligence data, where the objective is to ensure that the minimum possible number of people learn the identities of agents and other information that might compromise intelligence sources and methods. At present, the systems used to process such data typically have a 'compartmented mode' security policy, which is essentially Bell-LaPadula but with a large number of compartments (Top Secret X for codewords X) with limited numbers of people cleared for each codeword. This has the effect of stopping information flowing across as well as down; but it can be hard to administer (the US Navy a few years ago had over a million distinct classifications of data). The difference between BLP and the BMA model is that in the former, the assignment of access rights via codewords is centralised, while in the latter it is as decentralised as possible. We suspect that the BMA model would be cheaper and more effective. But the choice between central and distributed control is ultimately a cultural one.

4 Inference

It is not sufficient to remove patients' names and addresses from their records in order to make them anonymous. If a database allows sufficiently detailed queries, then individuals can still be identified, and this is especially so if information about different clinical episodes and even different individuals can be linked. For example, if I can make a query such as 'show me the records of all women aged 36 with daughters aged 14 and 16 such that the mother and exactly one daughter have eczema' then I can identify individuals. However, if I cannot make enquiries at this level of detail, then some legitimate research projects (such as looking for causes of cancers) could be made more difficult.

This problem was first seriously studied in the context of census information. A census collects a vast amount of sensitive information about individuals, and then makes statistical summaries of it available by geographical (and governmental) units such as regions, districts and parishes. This information is used not just in the general formulation of policy, but also in determining the levels of government funding to be made available to various local councils for the next ten years. So techniques for de-identifying data are important.

They are also becoming increasingly important in medicine, as numerous uses are developed for medical records outside the context of immediate care. Applications from research through drug marketing to health service cost control are increasingly making use of medical data, which often has to be de-identified in order to satisfy both ethical and data protection requirements.

There are two broad approaches to this problem, depending on whether the data are de-identified before or during processing – or equivalently whether the software that will process the data is trusted or not.

An example of the first kind of processing is given by the treatment of US census data until the 1960's. The procedure here was that one record in a thousand was made available on tape, minus the names, exact addresses and other sensitive data and with some noise added to the values that were given in order to prevent people with some extra knowledge (e.g., of the salaries paid by the employer in a company town) from tracing individuals. Totals or averages of variables such as people in various income groups were given, and records with extreme values (such as very high incomes) were suppressed. (Note that in order to prevent them being deduced from averages or totals, some other, non-sensitive data, must also be suppressed.)

In addition to the sample records, local averages were given. However, even this is not completely straightforward. Consider the situation of a wealthy family living in a small village: their income might make a significant difference to the per capita village income, and thus be deduced on the assumption that the per capita income of the other villagers is no different from that in nearby villages. So we might wish to suppress extreme values before the averaging process.

A more modern example is a system used for analysing trends in prescribing. Here, prescriptions are collected (minus patient names) from pharmacies; a further stage of de-identification removes the doctors' identities, and the information is then sold to drug company marketing departments. One problem with an early prototype of this system was that it merely replaced the names of doctors in a cell of four or five practices with 'doctor A', 'doctor B' and so on; an alert drug rep would notice, for example, "Well, doctor B must be Susan Jones because she went skiing in February and look at the fall-off in prescriptions here." The fix was to replace absolute numbers of prescriptions with the percentage of each doctor's prescribing which went on each particular drug. (This is a good example of the extent to which one must understand the application in order to design a system with appropriate inference controls.)

In the second type of processing, the data are retained in identifiable form on a database, and privacy protection comes from controls on the kind of queries that may be made. Early attempts at this were not very successful: famously, Dorothy Denning showed in 1979 that the US census was insecure. It was possible to construct a number of enquiries about samples containing a target individual, and work back to obtain supposedly confidential information about that individual. She demonstrated this by working out her boss's salary from legitimate enquiries on the census database.

If our census system allows a wide range of statistical queries, such as 'tell me the number of households headed by a man earning between $50,000 and $75,000', 'tell me the proportion of households headed by a man aged ($x - x + 5$ years) earning between $50,000 and $75,000', 'tell me the proportion of households headed by a man earning between $50,000 and $75,000 whose children have grown up and left home', etcetera, then one can quickly home in on an individual and infer his salary range. Such queries, in which we add additional circumstantial information in order to defeat the averaging, are known as trackers; they are usually easy to find. The obvious fix for this is query set overlap control (recording who has enquired about what) but this is rarely practical because of the amount of state that would have to be retained, and the ease with which users could collude or a single person acquire multiple accounts on the system.

One possibility is to control the type of enquiries that can be made. Maximum order control limits the number of attributes that any query can have: one study found that of 1000 medical records, three attributes were safe while with four attributes, one individual record could be found and with 10 attributes most records could be isolated. A more thorough approach (where it is feasible) is to reject queries that would partition the sample population into too many sets; in this way the extraction of individual records can be prevented directly.

The most common countermeasure is simply to specify a minimum query size. For example, New Zealand's National health Information System databases will reject statistical queries whose answers would be based on fewer than six patients' records. This is not enough in itself; if I can make an enquiry on six patients' records, and then on those records plus the target's, then I can break the system. Similarly, if there are N records in the database, it must not allow me to enquire on $N - 1$ of them, especially if global averages are known.

A more sophisticated approach is to make all the query sets the same size, often in conjunction with random sampling. This is now used by the US census: each query is answered by reference to a pseudorandomly selected sample of the actual data. This may be combined with adding random noise to it, sometimes by a system of rounding or ranging.

Some of the most powerful attacks come when the attacker can insert or delete records into the database. He can add records in order to make the target part of a group that is large enough for statistics to be revealed, and then check that he did indeed catch a single individual by deleting a single record and verifying that the query is now blocked. One countermeasure is insisting that new records be added or deleted only in batches. Taking this to an extreme gives is partitioning – in which records are added in groups and any query must be answered with respect to all of them or none. However, this is equivalent to publishing tables of microstatistics.

The quality of protection offered by a statistical database can be measured in a number of ways. The US census uses the 'n-respondent, k%-dominance rule': it will not release a statistic of which $k\%$ or more is contributed by n values or less. An alternative approach is 'precision': if we let D be the set of statistics that are disclosed and P the set which must be protected, then we need $D \subseteq P$ for privacy, and $D = P$ for data protection. If $D = P$ then the protection is said to be precise.

Finally, a related problem to inference is that an opponent who gets hold of a number of unclassified files might deduce sensitive information from them. For example, New Zealand journalist Nicky Hagar deduced the identities of many officers in GCSB (their equivalent of GCHQ) by examining lists of service personnel and looking for patterns of postings over time. Intelligence officers' cover postings might also be blown if an opponent gets hold of the internal phone book for the unit concerned and finds that the officer is not listed there. The army list might be public, and the phone book 'restricted'; but the fact that a given officer is involved in intelligence work might be 'secret'. Combining low level sources to draw a high level conclusion is known as an aggregation attack; it is clearly related to (but not the same as) the increased risk to personal information that arises when databases are aggregated together.

While inference attacks can be countered by statistical techniques, neither type of aggregation threat can. If possible, we might monitor who tries to access what; otherwise, it is likely that we will have to manage a residual risk.

5 Further Reading

Dorothy Denning was a pioneer of inference security, and a good description of the basics can be found in her book *'Cryptography and Data Security'*.

6 Problems

1. Given a system that implements the BIBA integrity model, what would you have to add to implement Clark-Wilson? Would the size of the TCB have to be increased?

2. You have been engaged to design an order entry and bookkeeping system which accepts orders by modem from a firm's sales representatives, has them verified by the sales manager, then generates shipping orders to the warehouse, invoices to the customer and monthly commission cheques to the salesmen. Credit limits are maintained on customers by a credit controller, and on salesmen by the sales manager, while the finance director needs accurate information on the overall cash, credit and stock positions. Describe how you would develop a security policy and mechanisms for such a system in the contact of a specific operating system (e.g., Unix).

3. A medical system replaces patients' names with postcode sector (e.g., CB2 3), year of birth, and a hospital serial number before the record are used for research. Assuming that there are 6,000 patients per postcode sector, what is the probability that a given individual will be identifiable to a researcher who does not have access to the index of hospital serial numbers?

4. You are preparing software for processing Tripos exam results. Only the examiners have access to actual marks; it is policy that these should not be published, but only the grade of pass. You are informed that the Old Schools now wish to publish correlation coefficients for each department between marks and candidates' country of domicile, type of school attended, sex and age. Provide an outline design of the mechanisms you would select to ensure that individual marks could not be inferred from the correlations.

5. The National Lottery keeps secret the amount of money wagered on each of the 49 numbers. However, the total amounts wagered and the prizes (including bonuses paid) are public. Assuming that the wagering pattern is constant over time, could you derive it from the public information?