

A Security Architecture to Protect Against Data Loss

Clive Blackwell

Information Security Group, Royal Holloway, University of London,
Egham, Surrey. TW20 0EX. UK
C.Blackwell@rhul.ac.uk

Abstract. Data loss poses a significant and increasing problem for organisations. This is shown by the regular stories of data loss reported daily in the media, such as the mailing of 2 CDs containing 25 million personal records by the Revenue and Customs in the UK. There is a need to provide systematic protection to data in all its forms and locations however it is accessed. We have developed Searchlight, a three-layer security architecture containing the physical, logical and social levels, which we use to analyse data loss holistically to prevent, detect and recover from exposure. We examine deliberate and accidental data loss by employees, but the same analysis can be straightforwardly applied to external attacks. Our practical security model appears to have widespread application to other problem domains such as critical infrastructure, the insider threat and financial systems, as it allows the analysis of systems in their entirety including human and physical factors, not just as technical systems.

Keywords: Data loss, security architecture, Searchlight model, attack and data loss classification.

System Modelling

Data Loss and Its Causes

We define data loss as the disclosure of sensitive confidential or private data held about an organisation or person, or metadata about systems and their means of access such as passwords. The effects of data disclosure on an organisation may be financial loss, interference with business activities, loss of trust and reputation, and stronger regulatory controls and fines. The ultimate victim may be a third party data subject whose information is held by the organisation, who may suffer financial and identity fraud, or loss of privacy or reputation, that may take time and effort to recover, and cause psychological distress.

We investigate data loss incidents caused by employees and other insiders with legitimate access, but the same analysis can be straightforwardly applied to external theft. Insiders can cause great damage to organisations because of their privileged access, knowledge of weaknesses and the location of valuable data that can be misused for their own purposes or inadvertently revealed. The insider threat is the biggest threat to organisations' intellectual property and other sensitive data according to 68% of respondents in a recent McAfee report [1]. The disclosure of the personal records of third

parties is also serious, as the cost of recovery averages \$202 per record, according to the Ponemon Institute [2]. Some new books cover data loss in more detail [3], [4].

Accidental Data Loss by Her Majesty's Revenue and Customs

We show how to use our systematic model to analyse data loss, by applying it to the loss of 25 million personal records by HMRC. The goal of the employee in the HMRC fiasco was helping the organisation satisfy its regulatory responsibilities. HMRC was required to reveal a specific subset of Child Benefit data to the National Audit Office for auditing. This led to the employee sending out all 25 million personal records by mail on 2 CDs that was an unacceptable breach of policy and lack of commonsense, which led to him losing his job, and the head of HMRC resigning. We use our model to analyse weaknesses in the data handling procedures that led to the incident, and determine the security measures that could have avoided it and minimise the likelihood of a reoccurrence.

The Searchlight Model

Multilevel Security Model

We believe that data loss is a complex multifaceted problem that requires systematic analysis to mitigate. We have designed a three-layer architectural security model called Searchlight to investigate and evaluate system security, which we apply to data loss. The use of layers is a common structuring method used to decompose and analyse systems. We are influenced by Neumann's practical classification system for attacks with eight layers [5], [6], which are, in descending order: the external environment, user, application, middleware, networking, operating system, hardware and internal environment.

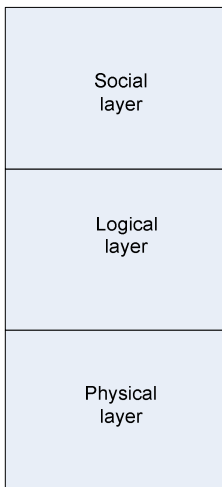


Fig. 1. The Searchlight model

The Searchlight model is a three-layer model, which includes the social layer (people and organisations) and physical layer along with the middle logical layer containing computers and networks. This allows a holistic representation and analysis of complex systems such as organisations in their entirety, including human and physical factors, rather than as technical systems alone.

Data has a separate spatial scope at each layer. For example, the number of people that know some information is its scope at the social level; the extent of logical data is its availability to computers and applications that may extend worldwide over the Internet; whereas physically stored data may be restricted to paper documents on a desk or in a filing cabinet.

The social or organisational layer contains the abstract representation of organisations by their attributes including their goals, policies and procedures. It also includes people and their characteristics such as their goals, knowledge and beliefs. Business and personal information exists at the social level, because it does not have meaning at lower levels. Data

may be stored, processed and transmitted on computers and paper at lower levels but this representation is not understood semantically. This is similar to Searle's argument with his Chinese room [7], where a program that simply transforms inputs of Chinese characters to corresponding outputs of Chinese outputs could possibly pass the Turing test without demonstrating any intelligence.

The logical layer is the intermediate layer that contains intangible computational entities including computers, networks, software and data. Logical data is a representation of information at the social level, where it is more amenable for processing, distribution and storage. It also includes data used to supply logical services to social level entities such as authentication information like passwords that link people to their accounts. The logical layer is incorrectly the focus of most attention in security, because all layers need protection to provide comprehensive security.

The physical layer is the bottom layer that contains tangible objects including buildings, equipment, paper documents, and the physical aspects of computers and associated devices. In addition, it contains electromagnetic radiation such as radio waves, electricity and magnetism that are used to transmit and store data. All higher layer entities, including people and information, have a physical existence as well as a higher layer representation that must be considered when analysing data security.

Examples of data loss from physical objects include insecure laptops, USB sticks and paper documents. Data can also be leaked in transmission over various physical communication channels with differing spatial scopes: overheard during face-to-face and phone conversations, and over computer communication channels such as the Internet, local networks, and from keyboard inputs.

Technical data protection measures alone are incomplete and cannot stop attacks that occur partially or totally at other layers. The comment that Butler Lampson and Roger Needham attributed to each other: "Whoever thinks his problems can be solved using cryptography, doesn't understand his problem and doesn't understand cryptography", can be extended to the inadequacy of technical controls in general. In addition, the social level controls such as policies and procedures can usually be evaded by employees, as they cannot cover every eventuality and are often weakly enforced. Physical attacks to steal, damage or misuse equipment, computers and documents are also common. We conclude that organisational security must involve all three layers to provide comprehensive defence against data loss.

An Attack Classification Scheme

We investigate the stages of data loss with an extension of Howard and Longstaff's taxonomy [5], [6] for network security incidents that show the different classes of entity involved and their relationships. Our classification scheme extends Howard's taxonomy to include the social and physical aspects of systems, which allows comprehensive modelling of organisations, including the corresponding defensive measures.

All incidents are initiated by people at the social layer and are only effective if they meet a social goal such as obtaining money, power, reputation or pleasure. However, people cannot operate directly at the logical layer, so they use agents to act on their behalf such as user accounts to issue commands, run programs and access services.

In the active stage of an incident, the *perpetrator* or their *agent* employs a *method* to perform an *action* that executes a *threat* to exploit a *vulnerability* with an

immediate effect on a *target*. This ultimately achieves the perpetrator's social layer goal, which could be benign and have an unintentional effect on the victim. We distinguish between the *immediate effect* at the lower layer on the confidentiality, integrity and availability of data loss, and the *ultimate effect* on the victim at the social layer of financial, privacy and reputational loss.

In addition, we have corresponding concepts to describe and classify defensive mechanisms, where there is a matching defensive category for each incident concept described above. For example, there are different defensive controls to stop the immediate effect of loss of data confidentiality using data loss prevention (DLP) tools, and to minimise the ultimate social level financial impact on a victim when their compromised credit card details are used in a fake transaction.

Possible incidents can be plotted in a table showing the active elements of our classification as columns in a grid with a row for each level. Incidents' progression through the various stages are shown as paths through the grid from left to right starting with access to the target before moving on to demonstrate the subsequent damaging effects on the organisation. We can then create an analogous table with the corresponding defensive measures, to provide a complete and consistent defence at all layers to prevent or constrain incidents, as we demonstrated elsewhere [10], [11].

Data Loss Classification

We need a clear understanding of the functions and weaknesses of data along with the powers of users that may be abused. We consider that sensitive organisational and personal information, and account access data such as passwords constitute the main targets for data disclosure. Impacts on reputation, recovery costs and psychological wellbeing may also be the primary motive of the data breach, but are generally side effects of the incident.

We can classify data loss incidents according to their effect on the victim or the benefit to the perpetrator. We can focus on the purpose of the perpetrator to classify their intent into financial gain, business advantage, psychological pleasure and accidental disclosure. Incidents can also be classified by their effects on the victim from damage, fraud and theft, which have an undesirable impact indirectly by first breaching the fundamental security service of confidentiality, usually at lower layers.

The effects on the organisation from the three classes of data loss are:

- **Damage** – Caused by reducing organisational reputation and business abilities. Indirect effects arise from interfering with its ability to perform its normal business activities by greater competition resulting from the disclosure of information about its business activities and products. Loss of reputation and trust arise from the release of sensitive third party information
- **Fraud** – Causes financial losses to the organisation or their customers by allowing unauthorised transactions, illegitimate access to resources, and impersonation to gain financial, business or personal services. An indirect attack caused by the loss of authentication or authorisation data
- **Theft** – Includes logical resources that give access to computer-held data, and physical assets such as computers, storage devices and paper documents

The impact to third party data subjects arising from an organisational breach includes:

- The loss of privacy from personal information lost, and possible interference with activities subsequently such as caused by identity fraud
- The time, effort and money to carry out recovery measures
- Psychological distress.

Attack Surface

Michael Howard, a Microsoft employee, invented the idea of the attack surface [1], which is the set of available channels to access and use computer systems. For example, it is the set of commands offered by an application or the available links on a Web page. We extend the idea of attack surface to all three layers, which allows a complete determination and analysis of exploitable access paths.

In addition, we extend the attack surface to include boundaries that the attacker can move through to gain local access to the target, rather than operating at a distance over a channel. Higher layers entities have a conceptual location that can describe their position, proximity and relationship to other entities at the same level. For example, every file has a logical position within a directory, and processes execute in virtual memory. People have a conceptual social-level position such as a role within an organisation, or a personal position that can be taken over by an identity thief.

A complete attack surface can provide systematic defence by constraining movement to and remote access of the target at every layer. Insiders are not limited by external system boundaries such as building entrances and firewalls that protect the organisation from external attack. An employee may instead be constrained by internal attack surfaces that partition the system with defensive controls, such as role-based access controls that must be breached to gain unauthorised access to sensitive data. Many insider incidents, however, use authorised access, such as accessing the target data using their own accounts, so there is no interposed attack surface. For example, finance or sales employees may use their permitted access to steal money from accounts.

Impact Zone

We also need to limit the scope and impact of data breaches, which includes limiting undesirable effects on the organisation and third parties. The impact zone is the set of resources affected that are unavailable, modified or disclosed illegitimately. TJ Maxx [13] allowed at least 45.7 million credit and debit card details to be revealed because of inadequate protection. Data disclosure can be limited by data minimisation as far as possible to reduce the impact of incidents.

The impact zone from disclosed data extends past the organisational or system boundary, so that access is no longer under its control. This is a dual notion to the attack surface that constrains the inward movement and access to organisational resources. The impact zone is already used informally in data protection at all three layers to limit the effects of disclosure. Employees are trained not to reveal sensitive information about the organisation to third parties on the phone. Data loss prevention (DLP) refers to controlling the disclosure of sensitive information by searching documents and messages for confidential information before release. Paper documents naturally have a restricted scope compared to digital data, as they are more

difficult to distribute physically. Finally, it applies to the rule of least privilege that limits employees' privileges to the minimum required for the job, which limits the impact of data incidents using authorised access.

The impact must have an ultimate effect at the social layer, as lower-level resources only have value to the extent that they support organisational goals. Data disclosure can only have an effect from its subsequent misuse. For example, the loss of sensitive business data, such as intellectual property and business plans, may lead to business disadvantage from stronger competition when it is used by their competitors.

The impact zone is fundamental to data disclosure, which must have an external scope to be effective. For example, fraudulent use of third party credit card details causes losses to third-party victims. To affect the organisation, however, the effects must boomerang back to the organisation by the actions of the victims who seek compensation or the authorities who levy fines. The effect of data disclosure is thus an external release through the impact zone that returns to the organisation in a transformed form from a third party.

Protecting against Data Loss

Accidental Data Loss by Her Majesty's Revenue and Customs

We apply the Searchlight model to analyse the loss of 25 million personal records by HMRC, as described in detail in the Poynter report [14]. An employee mailed all 25 million Child Benefit records on 2 CDs to the National Audit Office to satisfy their regulatory responsibilities to audit Child Benefit cases, which never arrived. This incident did not have a direct effect on HMRC, as the data concerned third party data subjects, but there were indirect effects on their reputation and trustworthiness, so they needed to improve their security procedures to minimise the likelihood of a reoccurrence. We indicate some of the main security measures to protect organisations, illustrated by the weaknesses of HMRC's data handling procedures, and refer the reader to [3] for an in-depth discussion.

System Hardening

The aim is to stop the ultimate social level effect on the organisation, so we can consider protective measures at multiple stages before, during and after a data exposure incident, which equate to attack surface reduction, hardening the target, and limiting the impact zone. The access and use of a data target is part of both the attack surface and impact zone, but is considered separately for clarity.

Sensitive data should be difficult to misuse or disclose, which requires comprehensive protection at all layers. There should be a complete attack surface to limit movement and routes to data at all layers to restrict unauthorised access and constrain authorised use. For example, the perpetrator of the HMRC data disclosure should not have been allowed access to the complete Child Benefit database.

The potential impact may be limited to the target, organisation wide or have an external impact on third parties. Redundant protection measures can provide defence-in-depth to provide multiple independent impact zones. This includes initial controls on the use of data internally, and subsequent controls on its export to other computers,

copying to storage devices such as CDs or USB sticks, or printing on paper. The HMRC disclosure would have been avoided had there been controls on copying data to CDs, or if the physical protection boundary was replaced with a logical boundary by encrypting the data before it left the organisation.

The impact may be limited to the lower physical and logical layers or cause an ultimate effect at the social layer. Additional protection measures may avoid interference with lower layer resources from harming an organisation's essential activities. For example, loss of credit card details can be overcome by spotting anomalous transactions or by compensating victims afterwards. Unfortunately, recovery from data exposure is very difficult, as data may take on a life of its own, especially on the Internet.

Targeting the Perpetrator

We showed in our attack taxonomy that all incidents are initiated at the social level by a person pursuing a goal. The actions may be well meaning if an employee accidentally reveals data when work tasks are incorrectly executed, as with the HMRC fiasco, or they may have dishonest motivation for personal gain. We use the perpetrator's goals to determine appropriate social-level protection measures to reduce attackers' motivation and increase honest employees' alertness.

Attacks are often prompted by the need to resolve or relieve personal and work problems. Personal issues include divorce, drug abuse, financial problems and emotional disturbance. Organisational issues include job dissatisfaction, workplace disputes and disciplinary sanctions. The organisation can reduce the threat by addressing employees' personal and financial issues, and encouraging greater loyalty by good work conditions, fair treatment and attending to grievances. They may deter attacks with strong defensive measures that make the cost/benefit equation less favourable by increasing the risk or reducing the benefits of stealing data.

We now indicate how to deal with accidental data exposure by employees. The perpetrator often lacks foresight of the possible consequences, as data is intangible and easy to undervalue. Employees' obligations should be made clear by the explicit allocation of duties and responsibilities, with well-publicised understandable policies enforced by disciplinary action for breaches, which increase their incentive to be careful because of the possible repercussions.

Accidental disclosure is often caused by employees trying to get their jobs done by solving problems and being helpful to others. Contributory factors include ignorance, poor morale, and a lack of loyalty or acceptance of organisational values and rules. There may be a lax corporate culture where data is inadequately valued with inconsistent procedures and inadequate policies, leading to a lack of care. Perverse incentives must be avoided, as in the failure of the financial system, where the reward structure pressures employees into risky behaviour. All of these factors were present in the HMRC data loss according to Poynter [14]. Accidental data disclosure by employees can be reduced by awareness and training to make them conscious of the possible risks, and there should be rewards for good behaviour.

System issues can be remediated by redesign to avoid so called 'accidents waiting to happen'. Organisations should make their systems easier to use correctly, and give visible indicators when sensitive data may be revealed. The complexity of the HMRC

systems contributed to the data loss, because of their lack of integration and usability, which caused a social-level issue because required tasks could not be easily executed.

Conclusions

Data loss poses a significant and increasing problem for organisations, because of its increasing quantities, uses and access routes. Systematic defence is required as no single method can protect against employees with legitimate access. We proposed an architectural three-layer security model to analyse complex systems. We provided attack and data classification models, and extended the attack surface with the dual concept of impact zone that allowed investigation of the stages in data loss. This enables a systematic determination of protection measures within the classes of limiting access to sensitive data, constraining the use of the data and limiting the impact of successful breaches. We also considered how to reduce the insider's motivation and accidental errors.

Our model has been used to analyse data loss incidents caused by employees and external attackers [3], but we focussed on accidental disclosure by employees, as illustrated by the HMRC fiasco. The corresponding determination of defensive measures helps to provide comprehensive protection against data loss incidents at all three layers from any source. It aids the provision of multiple supporting controls offering defence-in-depth, including recovery methods that limit the impact of incidents that are difficult to avoid.

Our security model appears to have widespread application in other areas such as the insider threat, critical infrastructure and financial systems, as it allows the analysis of systems in their entirety including human and physical factors, not just as technical systems. We have applied to model to the insider threat [10], [11] and to investigate critical infrastructure with its widespread scope and weaknesses at all layers [15]. In addition, it has application to complex financial systems, such as banking networks where weak procedural and physical controls are usually exploited, rather than the technical controls such as cryptography [16].

References

- [1] McAfee, Unsecured economies: protecting vital information (2009), <http://resources.mcafee.com/content/NAUnsecuredEconomiesReport>
- [2] Ponemon Institute. 2008 Annual Survey: Cost of a Data Breach (February 2009), http://www.encryptionreports.com/download/Ponemon_COB_2008_US_090201.pdf
- [3] Blackwell, C.: Data Loss: the Essentials (September 2009), IT Governance at <http://www.itgovernance.co.uk> or <http://www.27001.com>
- [4] Bunker, G., Fraser-King, G.: Data Leaks for Dummies. Wiley, Chichester (2009)
- [5] Neumann, P.G., Parker, D.: A Summary of Computer Misuse Techniques. In: Proceedings of the 12th National Computer Security Conference (1989)
- [6] Neumann, P.G.: Practical Architectures for Survivable Systems and Networks. SRI International (2000), <http://www.csl.sri.com/neumann/survivability.pdf>

- [7] Searle, J.R.: *Minds, Brains, and Programs*, from *The Behavioral and Brain Sciences*, vol. 3. Cambridge University Press, Cambridge (1980),
[http://web.archive.org/web/20071210043312/
http://members.aol.com/NeoNoetics/MindsBrainsPrograms.html](http://web.archive.org/web/20071210043312/http://members.aol.com/NeoNoetics/MindsBrainsPrograms.html)
- [8] Howard, J.D.: *An analysis of security incidents on the Internet 1989-1995*. Carnegie Mellon University (1997), <http://www.cert.org/archive/pdf/JHThesis.pdf>
- [9] Howard, J.D., Longstaff, T.A.: *A common language for computer security incidents*. Sandia National Laboratories (1998), <http://www.sandia.gov>
- [10] Blackwell, C.: *The insider threat: Combating the enemy within* (2009), *IT Governance* at <http://www.itgovernance.co.uk> or <http://www.27001.com>
- [11] Blackwell, C.: *A Security Architecture to Model Destructive Insider Attacks*. In: *8th European conference on information warfare*. Academic Publishing Ltd. (2009)
- [12] Howard, M.: *Attack surface: mitigate security risks by minimizing the code you expose to untrusted users*. *MSDN Magazine* (November 2004),
<http://msdn.microsoft.com/en-us/magazine/cc163882.aspx>
- [13] MSNBC. T.J. Maxx data theft worse than first reported (29 March 2007) MSNBC at: <http://www.msnbc.msn.com/id/17853440>
- [14] Poynter, K.: *Review of information security at HM Revenue and Customs*. HMSO (2008),
http://www.hm-treasury.gov.uk/d/poynter_review250608.pdf
- [15] Blackwell, C.: *A Multi-layered Security Architecture for Modelling Complex Systems*. In: *4th Cybersecurity Information Intelligence Research Workshop*. ACM Press, New York (2008)
- [16] Anderson, R.: *Why cryptosystems fail*. In: *1st ACM conference on computer and communications security*. ACM Press, New York (1993)