

Incident Response and Data Protection

Document Version: 02

Date: September 2011

Author: Andrew Cormack (JANET(UK))

Abstract

This paper discusses how the use of information by Computer Security Incident Response Teams (CSIRTs) is justified under European Data Protection legislation, specifically Directive 95/46/EC. CSIRTs need to use information from computers and log files to prevent, detect and respond to computer security incidents, however some of this information may constitute personal information as defined by the Directive and it may be impossible for CSIRTs to establish prior relationships with all those whose information they may handle, or even to notify them after the event. The Directive permits this kind of processing, provided certain conditions are met. The paper identifies these conditions and suggests measures that CSIRTs may use in planning and performing their activities to satisfy the requirements.

The paper is not intended as legal advice and should not be taken as such. No responsibility will be accepted by the author, his employer, or the publisher for any errors or inaccuracies, or for any consequences of acting on the contents of the paper.

[Version 2 of the paper adds new examples of dealing with denial of service attacks and automated processing of incident data, as well as an Annex on information sharing agreements]

Contents

Incident Response and Data Protection	1
1. Introduction	3
2. Purpose of Processing.....	4
3. Factors to be considered.....	5
3.1 Is Use/Disclosure Necessary?	6
3.2 Does the Action Support Legitimate Interests?.....	6
3.3 Are the Data Subject's Interests Protected?	6
4. Is processing justified?.....	8
5. Examples of Application.....	10
5.1 Malware analysis honeypot.....	10
5.2 Forensic data from compromised machine.....	12
5.3 Flow data for detecting compromised systems.....	15
Collection	15
Disclosure.....	16
Disclosure of Denial of Service Victims.....	17
5.4 Automatic Processing.....	18
Darknet Mesh.....	19
IP Address Threat Lists.....	20
6. Other requirements.....	22
6.1 Notifying the Individual.....	22
6.2 Notifying the Regulator	22
7. References.....	22
Annex: Information Sharing Agreements	23

1. Introduction

Computer Security Incident Response Teams (CSIRTs) work at the overlap of legally recognised human rights. On the one hand the role of CSIRTs is to provide secure networks and systems on which users can exercise their rights to communicate freely and, if they wish, privately. On the other the action necessary to achieve that security may itself infringe on the privacy of communications. This paper endeavours to help achieve the right balance between those, sometimes conflicting, rights.

As part of their job to prevent and investigate incidents on the Internet, Computer Security Incident Response Teams (CSIRTs) often handle information that is associated with identifiers such as Internet Protocol (IP) or e-mail addresses. For example a compromised computer used to send spam will often contain all the e-mail addresses to which spam was sent and the IP addresses of the hosts from which it came, while logs from firewalls and intrusion detection systems will normally be tagged with the IP addresses of the computers that may have been the sources and targets of hacking or denial of service attacks. To resolve incidents, CSIRTs need to use this information themselves and may also wish to disclose it to others, for example to inform individuals or banks of a phishing attack or to warn potential victims of a new virus threat.

Since most CSIRTs will only rarely be able to link these e-mail or IP addresses to an individual, it is not clear in law whether they in fact are personal data as defined by privacy legislation (for example the UK *Data Protection Act 1998* states that information is only personal data in the hands of someone who is likely to be able to link it to an individual). However the EU's Article 29 Working Party has stated (on p.17 of Opinion 4/2007 on the Concept of Personal Data) that if an identifier such as an IP address might be personal data (a CSIRT may occasionally collect an address whose owner it can identify) then it should be treated as if it is. The source of information is not relevant – personal information processed within the EC is subject to European data protection law no matter where or how it was collected. In any case information that relates to security incidents should be handled with care and not disclosed unnecessarily so following good practice for personal data is also likely to be good practice for limiting the harm caused by inappropriate disclosure of security information.

This paper therefore examines the implications of European privacy law – specifically the *Directive on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (95/46/EC)* – for CSIRTs handling information relating to incidents. In particular it examines when and how it is appropriate for a CSIRT to use information itself, and the circumstances in which it may be appropriate to disclose it to others.

CSIRTs are also likely to handle other kinds of personal data in ways that are familiar from other types of organisation, for example maintaining lists of people responsible for security in each organisation in their constituency or disclosing information to law enforcement, courts

or regulators as a legal requirement. Since these uses of personal data are not unique to CSIRTs they are not considered here.

2. Purpose of Processing

Both a CSIRT's own use of information and any disclosure to others constitute "processing" as defined by European law. Directive 95/46/EC provides a number of justifications for processing personal data, of which a number may be relevant to particular aspects of CSIRT work. However the most general justification is provided by Article 7(f):

"processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject which require protection under Article 1(1)."

Indeed Directive 2009/136/EC amending the Telecommunications Framework Directives specifically mentions Article 7(f) as covering "preventing unauthorised access to electronic communications networks and malicious code distribution and stopping "denial of service" attacks and damage to computer and electronic communication systems" (Recital 53).

The main purpose of most CSIRTs is to protect the legitimate interest of their particular constituency. Depending on the CSIRT, the constituency may be defined in many ways: users in the same organisation, users and organisations connected to a particular network or using a particular product, customers of the CSIRT, users and organisations in a particular country, etc.

As well as protecting its own community, the CSIRT may wish to inform other parties who are affected by the problem it has discovered, for example:

- Individual users whose passwords or credit cards details have been disclosed, or their computers compromised;
- Organisations that are being targeted by an attack (for example a bank that is the target of a phishing attack);
- An ISP who is connecting a compromised machine to the Internet;
- The Internet as a whole, if the CSIRT discovers a widespread vulnerability or attack.

Article 7(f) may justify all of these, however any particular use or disclosure must satisfy all the three tests in the article: that it is

- a) “necessary”
- b) in the “legitimate interests” of either the CSIRT or a third party, and
- c) that these interests are not “overridden by the ... fundamental rights and freedoms of the data subject”

These tests are considered in turn in the next section.

It should be noted that the information is rarely given deliberately to the CSIRT by the people and organisations it relates to. Even in the case of a compromised machine, the CSIRT may be given it by the owner of the machine, but it is likely to contain information relating to many other individuals. The CSIRT is more like someone who finds a file of information left in a public place, or who watches a static closed circuit television camera for indications of mis-behaviour. This distance from the individuals concerned may on the one hand make privacy breach less likely, but on the other it may also make it very hard to inform them directly of what has happened (see section on Other Requirements below). In some cases it may not even be possible to verify the origin or accuracy of the information (just like a paper file found in a public place). Especially in such situations CSIRTs should remember that the information may be inaccurate and ensure that their own actions and those of anyone to whom they disclose the information take this into account.

3. Factors to be considered

The “legitimate interests” justification provided by Article 7(f) is unique in requiring a balance to be struck between the interests of the individual whose information is being processed (known as the **data subject**) and those of the third parties. All the other justifications have simpler tests, which require only that the processing be necessary and legitimate. Rather than simple yes/no tests, the balance in Article 7(f) requires a CSIRT to assess the relative strengths of the arguments for and against acting in a particular way. How necessary (or desirable) is it to act and how strong is our legitimate interest in doing so? How great a risk does our action represent to the individual whose information is processed? Only if the necessity for a particular CSIRT to act is clearly greater than the risk to the individual may that CSIRT act.

This assessment is likely to vary from incident to incident, and even between different CSIRTs involved in the same incident, so it is impossible to give hard and fast rules. However the following sections provide some ideas on how to assess the strength of the arguments for and against a particular course of action, concentrating on each of the tests in the Directive: necessity, legitimate interests, rights of the data subject. In each case the aim is to produce a rough scale of need, legitimacy, or protection against which particular circumstances can be judged.

3.1 *Is Use/Disclosure Necessary?*

Both security and privacy indicate that a CSIRT should not use or disclose any information if this is unnecessary, but the requirement to strike a balance means that even actions that seem necessary must also be assessed to determine how strong the need for them is. Since, in most cases, the purpose of a CSIRT's processing will either be to reduce the impact of a current incident (for example informing a bank that a credit card number has been stolen) or to prevent or reduce future incidents (for example sharing information about machines that have been compromised and may in future be used for spamming or Denial of Service attacks), the severity of the potential incident is a useful measure of the importance of taking a particular course of action. In fact it is the likely reduction in severity that should be considered: no matter how severe an incident might be, if the proposed action does not reduce the impact of the incident then there is little argument for taking it.

In establishing a scale for incident severity we note that an incident that gives control of one or more computers can subsequently give rise to incidents of many different types and severities. This type of incident is therefore considered more serious than one whose impact is more specific and more easily contained, for example the loss of a credit card number or a Denial of Service (DoS) attack. Incident severity also relates to the number of computers, organisations or individuals likely to be affected. An incident, such as a new vulnerability in a widely used operating system, that could affect systems throughout the world, is therefore considered more serious than one that only affects a few organisations or individuals, such as the disclosure of a username and password. A scale for incident severity, and therefore strength of need to act, could therefore read **global compromise > local compromise/DoS > personal loss**

3.2 *Does the Action Support Legitimate Interests?*

Article 7(f) is clear that any action must support the legitimate interests either of the CSIRT or else of the organisation or individual to whom the information is disclosed. Acting in the legitimate interests of the CSIRT or its parent organisation therefore seems to provide the strongest justification for action; CSIRTs also represent their constituencies, so are likely to consider protecting the interests of constituency's interests as a legitimate aim. The CSIRT has a less strong relationship with those outside its constituency, but may still wish to disclose information to these where they are directly affected (e.g. a bank or user involved in a phishing attack), and this appears to be within the scope of Article 7(f). Finally a CSIRT may feel it has some justification for acting in the general interests of the Internet, since all constituents ultimately rely on it, but this is likely to be the weakest justification. A scale of "legitimacy of interest" in terms of who the processing will benefit might therefore be **self > constituency > recipient > Internet**.

3.3 *Are the Data Subject's Interests Protected?*

As described above, Article 7(f) states that protecting the fundamental rights of the individual whose information is being processed can override the need to perform processing or disclosure. Actions that provide better protection for those fundamental rights are therefore

more likely to be justified under Article 7(f). Three factors in particular can affect the individual's rights: what identifier, if any, is associated with the information being processed or disclosed; whether the information was gathered by a process that limits the information gathered (e.g. a firewall log with defined fields) or one with no built-in limits (e.g. a forensic backup of a compromised computer); and how widely and to whom information will be disclosed.

As noted above, although the information handled by CSIRTs is generally associated with identifiers such as usernames, Internet Protocol (IP) and e-mail addresses, most CSIRTs will not have access to the information required to link these identifiers to living individuals. Where an identifier belongs to a member of the CSIRT's constituency there is, perhaps, a slightly higher likelihood than when an identifier comes from a different organisation, network or country. CSIRTs may therefore regard processing an in-constituency identifier as presenting more of a risk to privacy than an external one (note, however, that disclosing an external identifier may give it to someone who can link it to an individual so such disclosure should only be done when the CSIRT is confident that the recipient will use the information properly). Privacy protection may also be improved by removing parts of an identifier that could refer to a single individual, for example by deleting the user part of an e-mail address or aggregating together in a single figure all the traffic from a block of network addresses. This may not always be possible, for example if the purpose of an investigation is to identify a particular compromised computer or user account. However in some cases it may even be possible to achieve the purpose while removing all potentially identifying information, for example when warning others of a new virus or vulnerability. A scale of privacy protection by identifiers could read **constituency** < **external** < **aggregated** < **none**.

How the information came into the CSIRT's hands may also affect the likelihood that using or disclosing it may cause a breach of privacy. Some types of information collection can be planned in advance, with the systems and processes being designed to collect some types of information and discard others. For example firewalls can be configured to log only packet headers, or intrusion detection systems to record only the first N characters of message content. With these types of system it is also possible to document what information will be collected and in what circumstances. However CSIRTs also have to deal with incidents where they do not control what information comes into their care, nor can they document this in advance. For example if a computer is compromised and passed to the CSIRT for investigation then it will contain whatever its users (and potentially the attacker) put there. This may well include commercially sensitive and personal files and e-mails, passwords and credit card numbers and even illegal material, since these are often stored on, and distributed through, compromised servers. Since the CSIRT cannot control what information it receives in this kind of **unplanned** information recovery it should probably be more cautious about using and disclosing it than information that was gathered in a **pre-planned** way.

If a CSIRT is considering disclosing the information to others, then the extent of that disclosure is clearly a factor in any privacy breach, as is the confidence the CSIRT has that the recipient of the information will not misuse the information. If it were possible to inform an individual directly that their information or computer had been misused then this would be

ideal, however it will be rare for CSIRTs to be able to identify a single individual. Often the best that can be done is to inform the person or organisation (sometimes another trusted CSIRT) that is responsible for the internet sub-domain, range of internet addresses, e-mail service, network or country where the affected individual appears to be. Such disclosure should be limited as far as possible while still giving reasonable assurance that the warning will be passed on. Where an attack relates to a particular service, for example a bank that is the target of a phishing attack, then it may well be justified to provide information to them. Information may indicate that there is a wider threat affecting the CSIRT's own constituency or other parts of the Internet. Information about such widespread threats should normally be disclosed within closed, trusted communities such as the Forum of Incident Response and Security Teams (FIRST) or the anti-virus or DNS communities, which have informal or formal agreements on how shared information will be used (examples of these agreements are in an Annex to this paper); although desirable, it may not always be possible to remove identifiers from the information, for example if warning that a particular server is disseminating malware. Finally there may be occasions when a threat is so widespread and serious that information has to be disclosed publicly. Doing so is likely to risk serious breaches of security (and privacy if any identifiers are attached) so this should only be considered if the alternatives are even worse. A scale of privacy protection in disclosure could read **world < trusted community < affected service < responsible organisation < user < none**

CSIRTs should apply the same test to disclosing information within their own organisation. In the interests of both security and privacy, information should only be disclosed when it is necessary to do so and where the necessity is not overridden by the interests of the individuals whose information is concerned.

4. Is processing justified?

Comparing the assessment of the necessity and legitimacy of the proposed action with the assessment of the harm that may be caused to an individual's fundamental rights, it should be possible to determine whether or not Article 7(f) allows the action to proceed. One way to present the assessments is to display the various scales alongside each other.

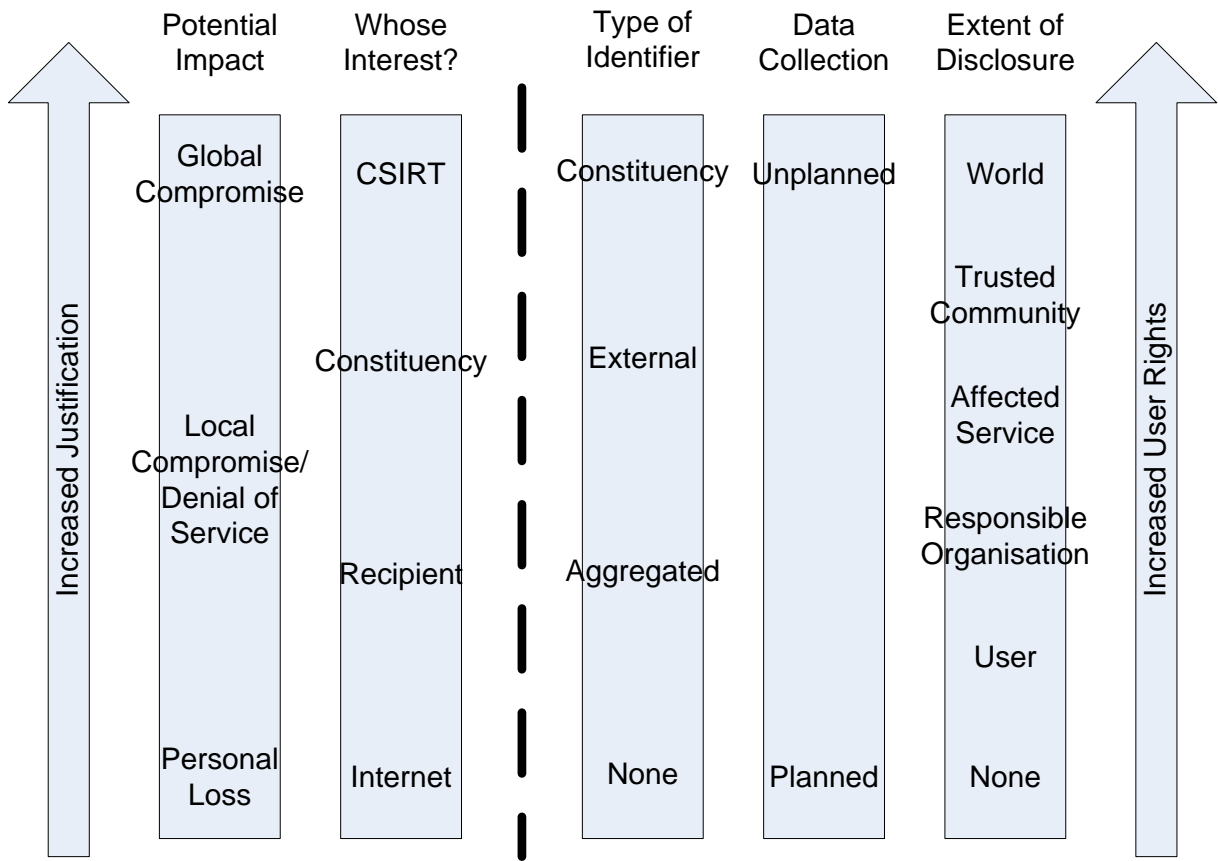


Figure 1 – Balancing Necessity and Risk

As in the examples in the next section, this visual presentation should make it apparent whether the proposed action is necessary for a legitimate purpose, or whether this purpose is overridden by the need to protect the fundamental rights of the individual.

5. Examples of Application

The following examples illustrate how this analysis can be used to justify and, sometimes, improve CSIRTs' activities.

5.1 Malware analysis honeypot

A number of projects have developed honeypots to capture various types of malware for analysis either by the honeypot operators or by anti-virus companies. The information can then be used to improve defences against malware.

Malware captured by such a honeypot is likely to represent a significant threat to a wide range of Internet-connected computers, which may well include those within the CSIRT's own constituency. In most cases the aim is to understand the malware itself so it may not be necessary to even collect potentially identifying information such as the source address; even if it is collected by the original honeypot it may not need to be included in any sharing of the resulting information. Honeypots can be designed to select the information they collect and log (for example by placing them on otherwise unused IP addresses so they should not receive any genuine traffic). Information about the malware will usually be shared within a trusted community of either honeypot operators or malware analysts.

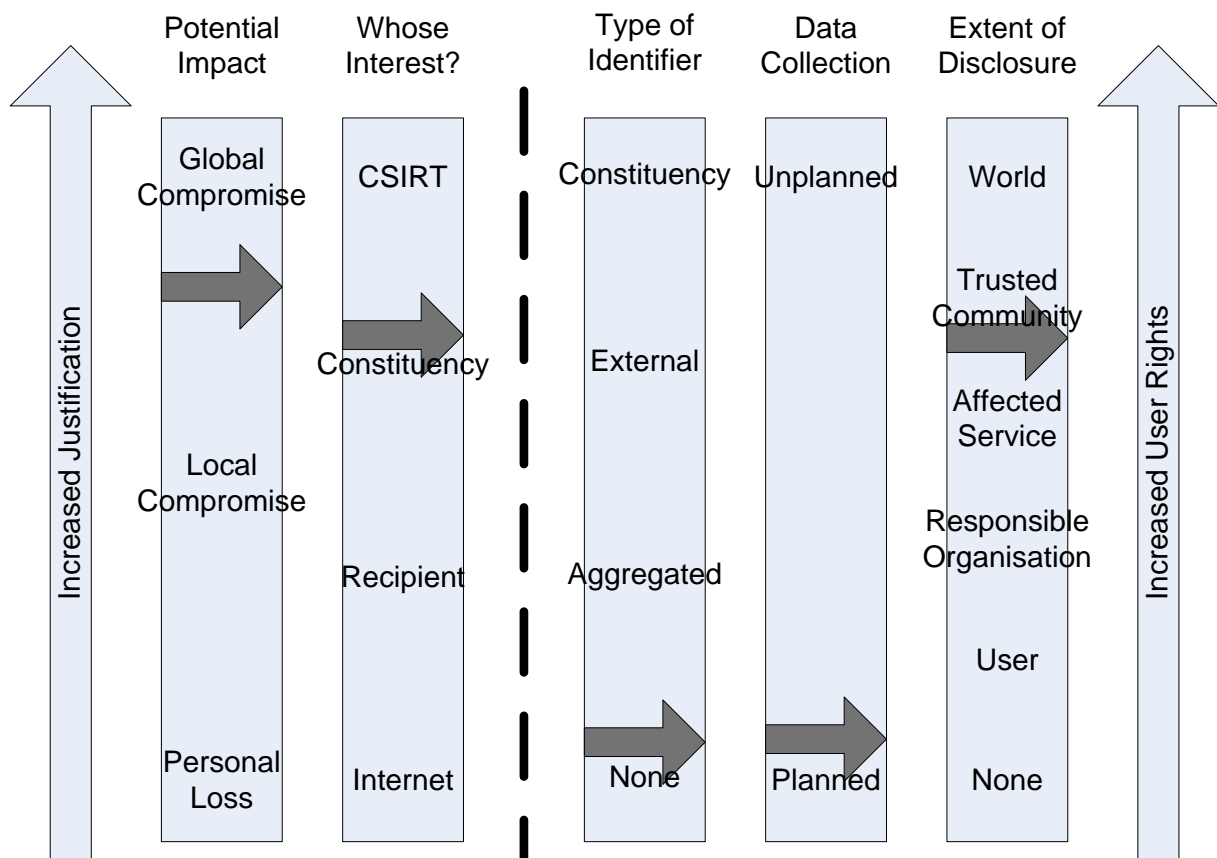


Figure 2 – Malware Honeypot

From the diagram it should be clear that the factors justifying the action (left-hand side) are stronger than those affecting user rights (right-hand side). It should also be apparent that this decision would be less clear if identifying information, such as complete IP addresses, were associated with the malware: in that case it might be necessary to reduce one of the other factors affecting user privacy, for example by disclosing only to the organisation responsible for each IP address.

5.2 Forensic data from compromised machine

A more complex example is where an incident response team has been provided with a copy of the disk from a compromised computer. Most CSIRTs will wish to identify how the machine was compromised and what it was then used for, then take measures to reduce the threat to their own constituency, then perhaps provide information to others who may have been affected. These different activities need to be considered separately under Article 7(f). Throughout it must be borne in mind that the CSIRT was not able to control the processes by which personal information came to be on the computer – both in legitimate use before it was compromised and unauthorised use thereafter – so there is no opportunity to design in protection of users’ rights before the event. Other measures will therefore be required to protect the rights of any individuals whose information may be present.

The initial investigation stage will normally be performed within the CSIRT and involve no disclosure of information. This stage may involve processing of all types of identifiers, however this higher risk to users will generally be balanced by the fact that the CSIRT is investigating a known compromise directly affecting its constituency.

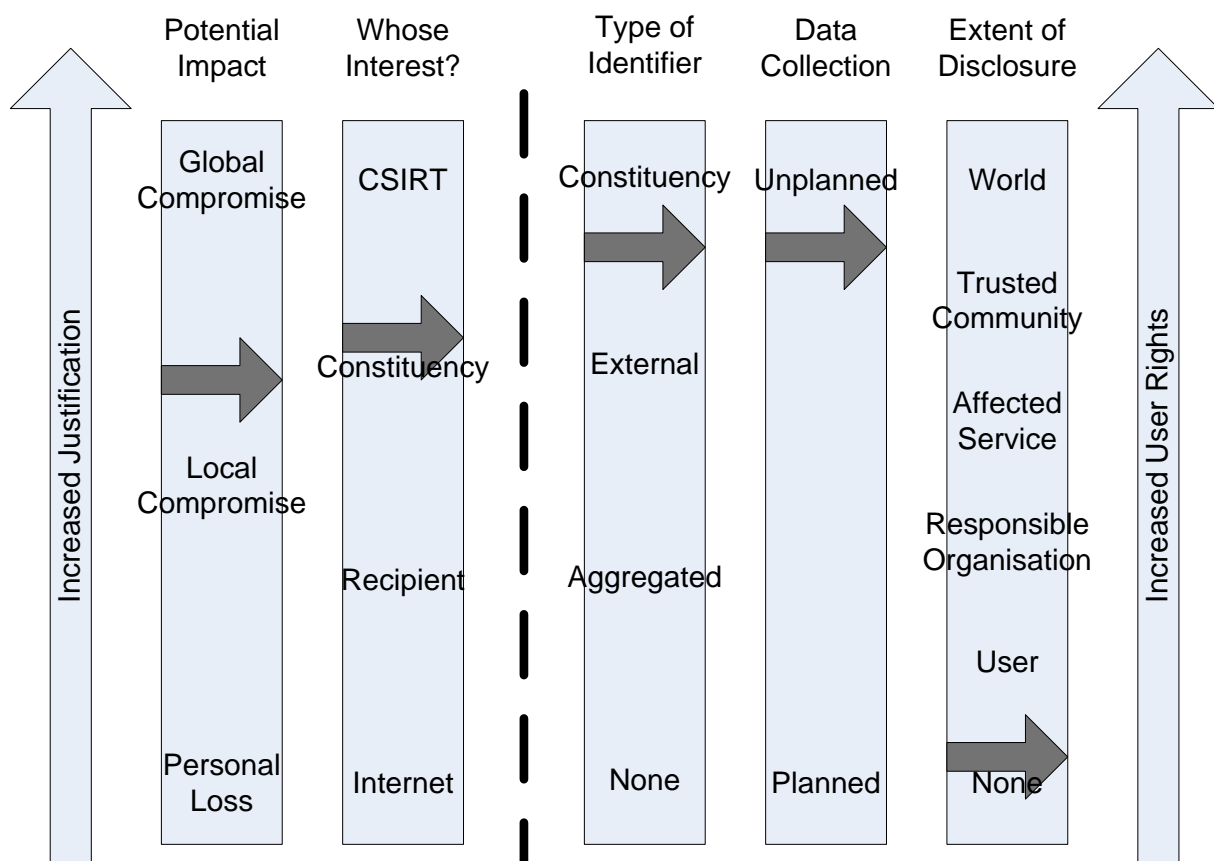


Figure 3 – Investigation of Compromised Computer

Reducing the threat to the CSIRT’s constituency will normally require some disclosure of information to members of that constituency. In order to keep the risk to user rights sufficiently low that it does not override the justification for the activity, the extent of

disclosure and the type of identifiers disclosed must be kept in balance: the dark grey arrows in Figure 4 below suggest that disclosure to the whole constituency of the full identifiers is likely to represent too much of a risk, whereas the light grey arrows indicate that the risk may be reduced to an acceptable level either by informing a trusted community of security contacts which members, but not which computers, (i.e. using a less intrusive Type of Identifier) were affected and inviting them to get in touch, or else by releasing to individual contacts the full identifiers of only the computers within their particular responsibility that were affected (i.e. reducing the Extent of Disclosure to only the Responsible Organisation).

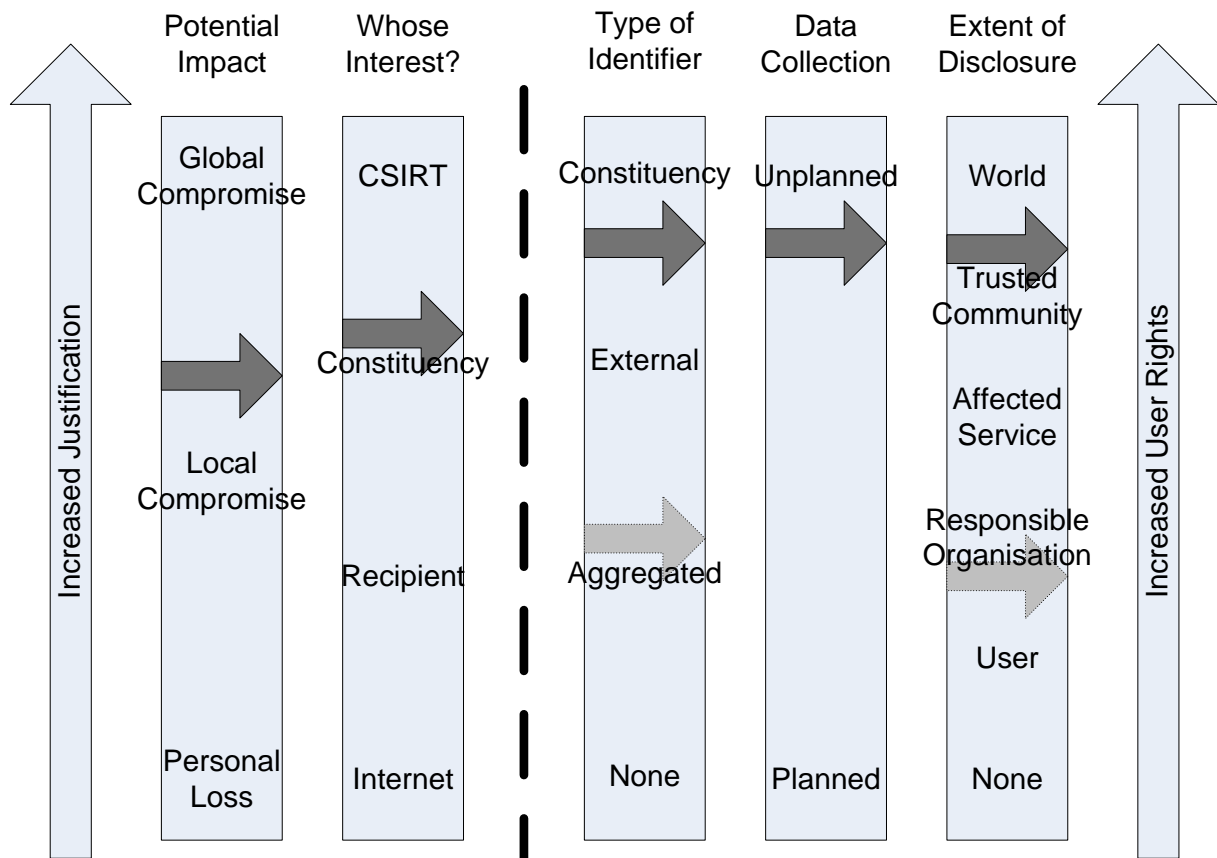


Figure 4 – Notification to Constituency

When considering informing those affected outside the constituency, the CSIRT has less justification for acting, but also less risk to privacy since the CSIRT will be dealing with external information that is less likely to allow it to identify the individual. To ensure that the justification for acting is not overridden by the risk to user rights, the CSIRT will need to choose very carefully the external organisations or individuals to whom the information is disclosed, since they may have additional information that allows them to identify (and potentially harm) the individual. It is therefore likely that disclosures will be limited to selected information sent to the services, organisations or users directly affected, since these have a strong interest in knowing about and resolving the problem. For example if a compromised computer is found to contain credit card numbers as a result of a phishing attack, those should probably only be disclosed to the relevant card issuers (i.e. Responsible

Organisation); user accounts for which passwords have been collected may be disclosed to the relevant users or organisations or, perhaps, to the operators of the services for which the accounts are used, though the light grey arrows indicate that the last of these (disclosure to the Affected Service) may be a borderline decision that may be hard to justify for lower levels of Potential Impact.

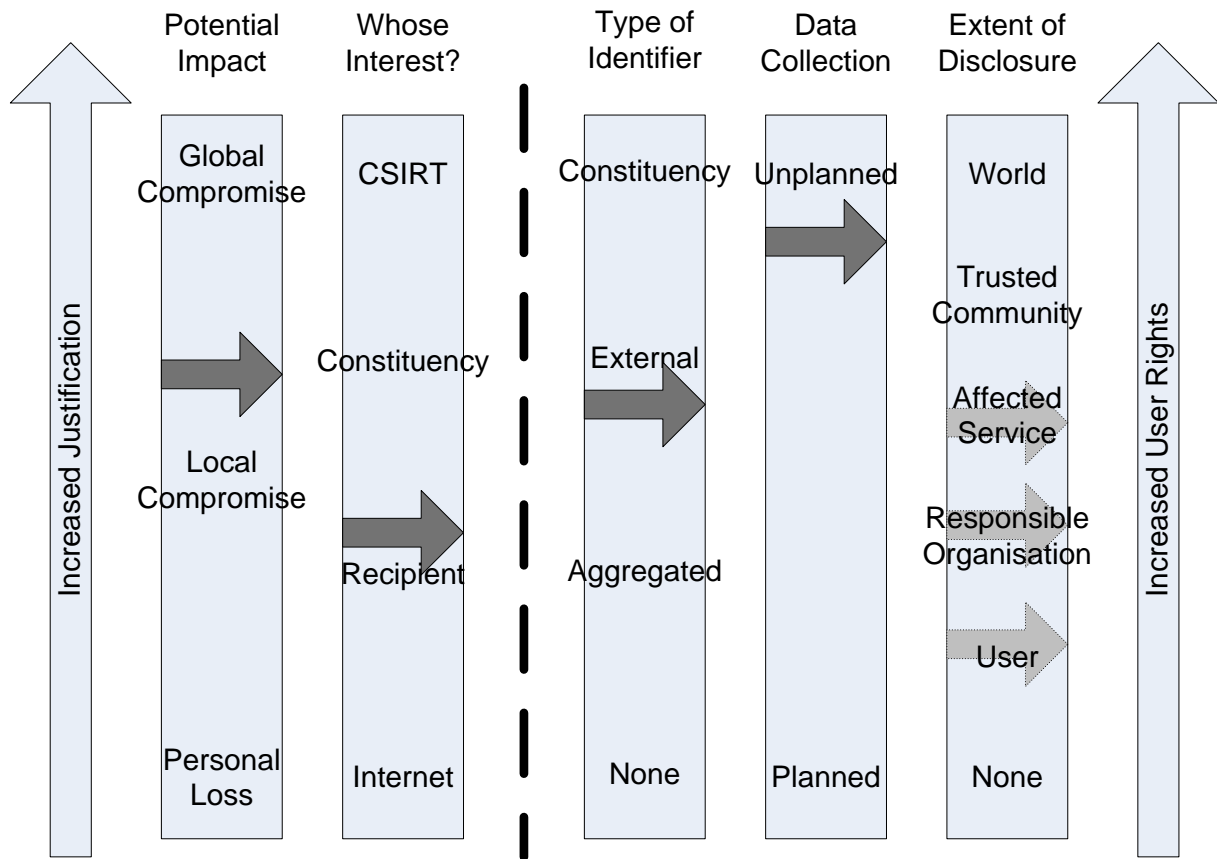


Figure 5 – Notification Outside Constituency

5.3 Flow data for detecting compromised systems

Collection

It is common for CSIRTs to collect information about network flows and use this to identify machines showing unusual patterns of behaviour that may indicate that they have been compromised. Flow data can, for example, be used to detect network scans, sources of spam or denial of service, or communications relating to BotNets. Depending on the particular circumstances, flow data may indicate that one of the machines in a flow represents a security threat (e.g. the source of spam, but not necessarily the recipients), or both (e.g. a BotNet controller communicating with one of its slave machines).

A number of different systems are in use, but most inspect the headers of Internet Protocol packets passing through a router and attempt to summarise the flows of communication represented by sequence of packets. Typical flow data will consist of a series of records each comprising a pair of IP addresses and the ports on which they are communicating, together with additional information on the number of packets and bytes involved and status information such as the direction of establishment of the flow, whether it started and completed normally, and any unusual characteristics, such as divergences from standard packet order. Some network flow systems will also capture some of the content of packets – these clearly represent a greater threat to privacy so need a different analysis from what follows.

Network flow records contain full IP addresses so do represent some risk to privacy. Addresses may be either internal or external to the CSIRT's constituency, though usually at least one of the communicating machines will be within the constituency. However the collection of flow records is planned and the information is used to detect compromised systems within the constituency: potentially serious security incidents in which the CSIRT has a strong legitimate interest. As the diagram shows, collecting and using network flow data for this purpose should therefore be acceptable within European privacy law.

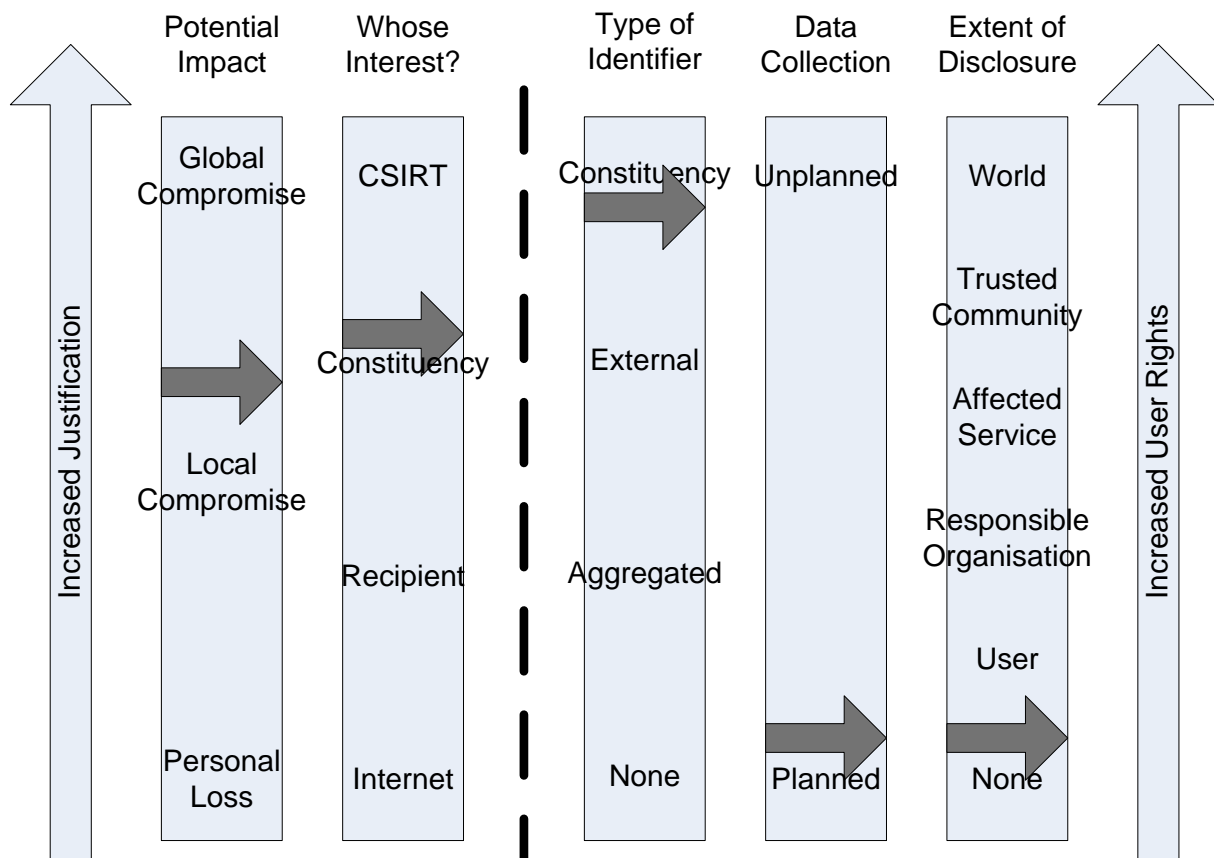


Figure 6 – Collection of Network Flow Data

Disclosure

As well as indicating security threats within the CSIRT's constituency, flow data may well provide an indication of problems elsewhere on the Internet. CSIRTs may well wish to disclose this information to those able to use it to deal with the problem at its source. Here there is a little less justification for the CSIRT's involvement, as the threat is external rather than within the CSIRT's own constituency, though clearly it is likely to be in their constituency's interest to have threats removed wherever possible. Fortunately the purpose of the disclosure is unlikely to be harmed by limiting disclosure in two important, and privacy-protecting, ways. First, information should only be disclosed to those likely to be able to deal with individual threat sources, for example the CSIRT or ISP in whose constituency the threat lies. Second, the recipient organisation should only need to know the full IP address of the threat source, not those of the corresponding machines within the disclosing CSIRT's constituency. These local addresses can either be removed from the flow data or anonymised by hashing or removing trailing component(s) of the IP address. This means that the disclosed information will only contain external identifiers, not those within the disclosing CSIRT's constituency.

As in the diagram these measures will normally reduce the risk to fundamental interests sufficiently to balance the reduction in the CSIRT's direct interest, thereby ensuring that this reduced disclosure is still justified.

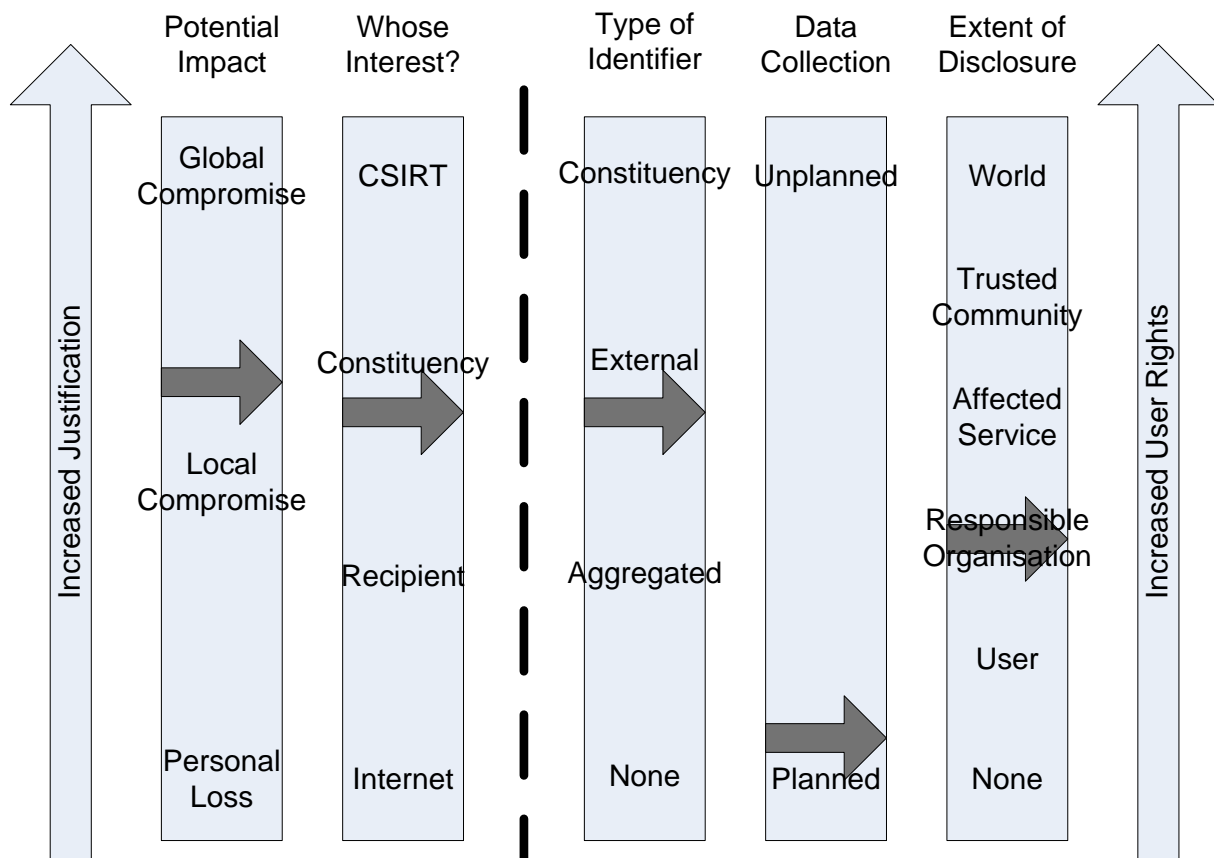


Figure 7 – Disclosure of Network Flow Data

Disclosure of Denial of Service Victims

Where a member of the CSIRT’s constituency is the victim of a Denial of Service (DoS) attack, much of the preceding analysis still applies. There still appears to be sufficient justification (as in Fig.6) for the CSIRT to collect flow data relating to the attack to protect both the network and the victim organisation; since DoS attacks are normally launched from compromised computers there is also a good case (as in Fig.7) for reporting to other CSIRTs or ISPs when machines in their constituencies are involved in the attack. However the CSIRT may also wish to ask others to help in mitigating the attack by filtering or blocking the attack traffic and here the analysis highlights a problem.

Where the purpose of disclosing flow data to others is to allow them to deal with problems in their own constituencies, only identifiers local to those constituencies (i.e. external to the disclosing CSIRT) need to be released. Furthermore those identifiers are each released only to the organisation responsible for them. If, instead, a CSIRT wishes to ask others to help reduce traffic to one of its own constituents then this requires disclosure of a constituency identifier – the IP address(es) of the victim – and this disclosure needs to be done to all the CSIRTs or ISPs from whose constituencies the attack is coming. As Fig.8 demonstrates, even if these teams are part of a Trusted Community, these differences may be sufficient to change the balance and indicate that the CSIRT should not act in this way.

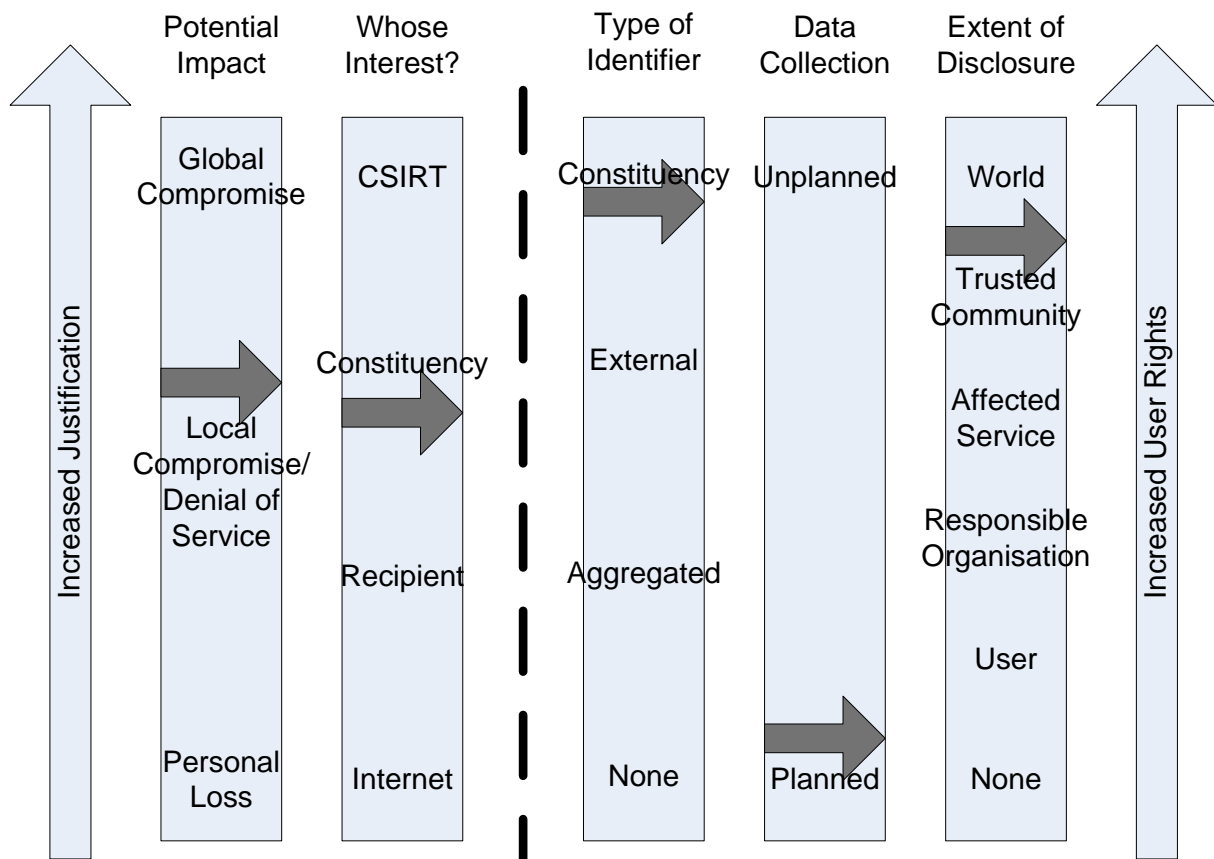


Figure 8 - Disclosure of DoS Victim

Although the CSIRT’s justification for acting, on the left hand side of the diagram, is unchanged from the previous example, the increased disclosure of more privacy-invasive information on the right hand side indicates that this disclosure may represent too much of a risk to the victim’s interests. Therefore a CSIRT in this position should probably not rely on Article 7(f) as justification for the disclosure, but should first discuss with the victim which course of action should be taken. For some organisations merely the fact that they are suffering under a denial of service attack may be considered sensitive information: these are likely to prefer to deal with the problem locally, with such assistance as the CSIRT can itself provide, rather than have the problem disclosed more widely.

5.4 Automatic Processing

The increasing amount of information that CSIRTs have to process and the increased complexity of many incidents has produced a growing interest in automating some of the processing and sharing of information. There is a well-established understanding that using computers, rather than humans, to process personal data can improve privacy. Computers instantly “forget” what they have seen so using them to filter information that is not relevant to an incident can avoid breaching the privacy of innocent bystanders whose traffic or data may have been captured along with that relating to the incident.

Automatic sharing of information might be seen as more of a threat to privacy, on the grounds that there is “no control” of individual disclosures such as would be provided by a human authorising each one. However if privacy is implemented through rules on what information may be shared and when, an automated system is actually more likely than a human to implement those rules accurately and consistently. Clearly such rules need to be designed very carefully as they will be the only thing protecting privacy; however automation means that rules can be designed and built as part of the controlled system development process rather than relying on an individual remembering, in the heat of an incident, to apply privacy checks before sharing.

The following sections give examples of how carefully-designed automated systems can both provide robust protection for privacy and provide rapid and effective incident response.

Darknet Mesh

Many organisations run darknet sensors (sometimes known as network telescopes). These are computers connected to the network but given an IP address or addresses that are otherwise unused. Since there are no services advertised by these systems and no users on them, they should receive no network traffic. All traffic that is received by the darknet sensor is therefore an indication of either a mis-configuration or suspicious activity by the computer sending it: for example it may result from a worm, virus or human scanning the IP address range in search of vulnerabilities. Darknet sensors can be very useful to the organisation operating them, since they allow the identification of suspicious and hostile incoming traffic that may represent a threat to the organisation’s other systems but that would otherwise be lost among the legitimate traffic handled by those systems. The collection of data by the organisation operating the darknet sensor is therefore justified in the same way as for flow data above in Fig.6 above.

However information from the sensor may also be useful to the organisation from which the traffic originates. A computer that sends packets to an unused IP address is at best mis-configured and at worst compromised. Such a computer may therefore represent a significant threat to users and data in its neighbourhood, and to the organisation within which it is located. Occasionally an organisation’s darknet sensor may capture packets originating on the local network in which case that organisation can also investigate the machine that is the source of the packets. To increase the likelihood of identifying local problems a group of universities have therefore agreed to share information from their darknet sensors, as described at <http://projects.oucs.ox.ac.uk/darknet/>. As well as using the information to protect their own networks against external threats, these universities each check whether traffic to their sensor originates from one of the other universities in the group. If any such packets are identified, an alert containing information about the relevant network flow(s) is automatically sent to the source organisation. This organisation then knows it needs to investigate a problem with the computer from which the traffic came. The privacy balance for this disclosure is similar to that for disclosure to a CSIRT’s constituency in Fig. 7 above, though the fact that the recipient organisation is not formally part of the discloser’s “constituency” might be considered to reduce the discloser’s justification for acting. However since the

agreements are reciprocal the discloser also benefits from increased information about its own network coming from sensors operated by others in the mesh, which gives it a stronger justification for participating. Concerning privacy, it is suggested above that automated processing according to pre-defined rules may present an even lower privacy threat than manual processing according to those same rules. The darknet mesh therefore appears to achieve a satisfactory balance between justification and privacy protection to be permitted under the Data Protection Directive.

[Note that in the case of a darknet sensor, there is no privacy issue in disclosing the IP address of the sensor, since it has no users. The operating organisation may therefore choose whether or not to obscure that IP address without affecting the privacy analysis above. There may be an operational issue, in that if darknet addresses become well known malware may avoid scanning them. However obscuring the IP address in the flow data does not avoid this problem, since the organisation receiving the flow data can compare it with its own logs and determine the IP address anyway! If an organisation is concerned about its darknet address(es) becoming known, it needs to agree with the source organisations it sends alerts to that they will keep this information secret.]

IP Address Threat Lists

A number of organisations use sensors such as the darknets discussed above to identify large-scale attempts to compromise other systems. For example, as described at <http://www.dragonresearchgroup.org/insight/sshpwauth-tac.html>, the Secure Shell (SSH) protocol provides significant security benefits by encrypting passwords when logging in to computers, but SSH servers often suffer from automated attacks that guess thousands of usernames and passwords in an attempt to break in to the servers and the systems they protect. In many cases a server that is compromised in this way will itself start to participate in future attacks, which are themselves harder to deal with because of the SSH encryption. These SSH Brute Force attempts are now seen daily on the Internet.

Network operators and system administrators across the Internet can significantly reduce the threat to their users and systems from brute force attacks if they have a list of the IP addresses that are their current source (typically there may be a few hundred addresses participating in attacks at any one time). Attempts to connect using SSH from these addresses can be either blocked or subject to additional checks. Since the attack sources change frequently, both the collection of information about attacks and the distribution of the resulting threat lists are most effective if they can be automated. As Fig. 9 below demonstrates, the benefit from publishing the threat list only justifies the privacy invasion involved in releasing to the world a list of unmodified IP addresses; automation, by ensuring that the information collection and sharing take place strictly as designed, provides a useful safeguard to ensure the correct balance is maintained.

Fig. 9 also suggests that a public threat list is only likely to be justified where there is a widespread threat of compromise affecting many systems. Furthermore, should a machine within the CSIRT's own constituency be listed this could tip the balance against disclosure,

so in this case the CSIRT must deal with the problem (and remove the IP address from the list) as soon as possible.

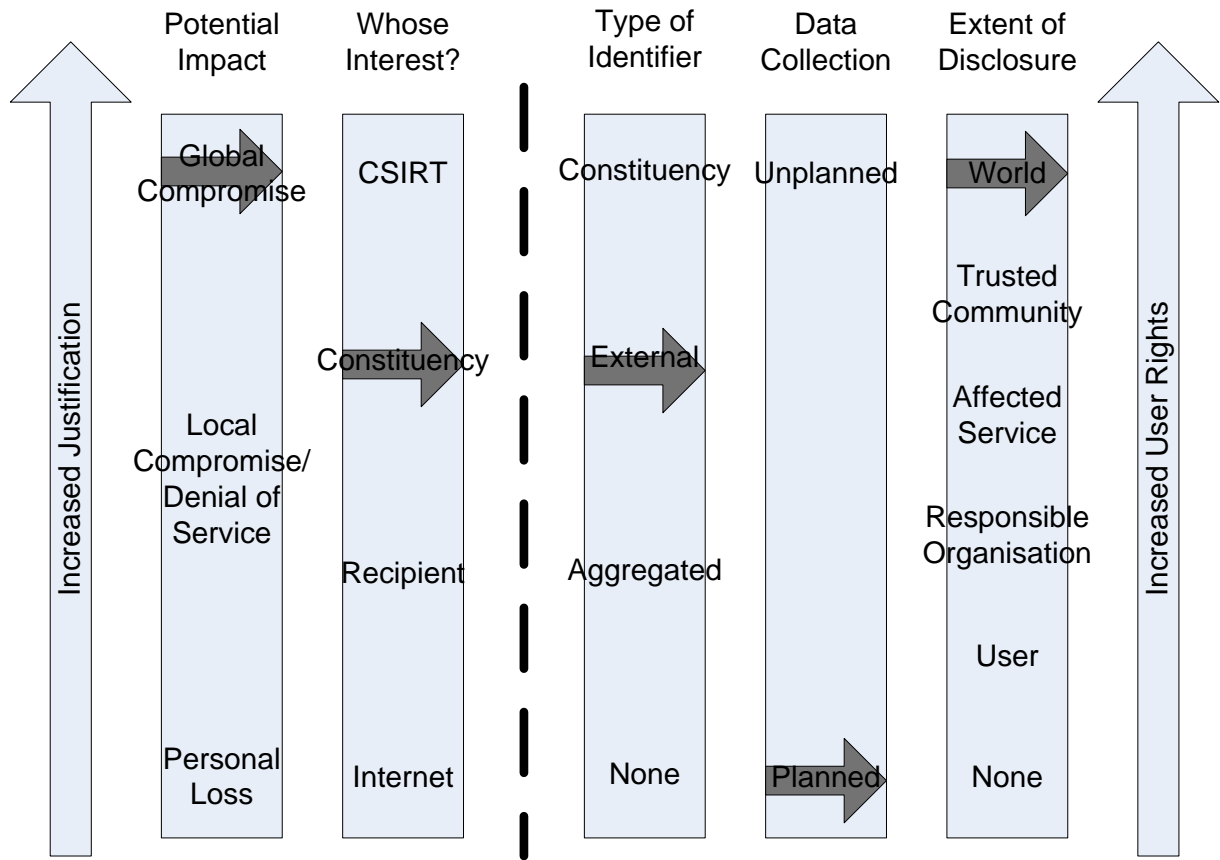


Figure 9 - IP Address Threat Lists

6. Other requirements

Directive 95/46/EC contains two additional requirements on notification that CSIRTs should be aware of, though compliance with them should raise few problems.

6.1 *Notifying the Individual*

Article 11 of the Directive applies where an organisation receives personal data other than from the individual. As discussed above, CSIRTs rarely receive incident data direct from the individual concerned, so if the information needs to be treated as personal data then they may be in this situation. Normally there is a requirement under Article 11(1) to inform the data subject of processing as soon as either recording or disclosure takes place. This would create a paradox if the information is not sufficient to identify the individual or to communicate with them (even if an IP address can be associated with an individual it cannot be used to send them an e-mail), however Article 11(2) allows the requirement to inform the individual to be waived in circumstances where “the provision of such information proves impossible or would involve a disproportionate effort”. In most cases where the information is sufficient to locate an individual person or machine, the CSIRT will normally inform either the person or a trusted contact for their network of the incident that has resulted in their information coming into the CSIRT’s hands. If this is not done then CSIRTs should think carefully about retaining the information since doing so may breach the Directive.

6.2 *Notifying the Regulator*

Article 18 requires that anyone processing personal data other than within a limited set of exemptions must notify their national regulator. Details of how to do this will vary depending on different national implementations of the Directive, however CSIRTs should normally be able to include their activities in either their own, or their parent organisation’s, notification.

7. References

Data Protection Directive (95/46/EC)

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML>

Amending Directive (2009/136/EC) for the Telecommunications Framework Directives

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:337:0011:01:EN:HTML>

UK *Data Protection Act 1998*

http://www.opsi.gov.uk/acts/acts1998/ukpga_19980029_en_1

Article 29 Working Party Opinion 4/2007 on the Concept of Personal Data

http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf

Annex: Information Sharing Agreements

The following are some examples of the types of agreement that may be used within trusted communities:

Forum of Incident Response and Security Teams

(<http://www.first.org/about/policies/op-framework/index.html>)

“All FIRST participants must adhere to the dissemination constraints specified by the originating source. Only the originator may relax any dissemination constraints. Information that has no specific dissemination instructions may not be disseminated further.”

Trusted Introducer for CERTs in Europe

(<https://www.trusted-introducer.org/links/ISTLP-v1.1-approved.pdf>)

Accredited teams use an Information Sharing Traffic Light Protocol, based on that developed by NISCC:

“RED Non-disclosable Information and restricted to representatives participating in the Information Exchange themselves only. Representatives must not disseminate the information outside of the Exchange. RED information may be discussed during an Exchange, where all representatives participating have signed up to these rules. Guests & others such as visiting speakers who are not full members of the Exchange will be required to leave before such information is discussed.

AMBER Limited Disclosure and restricted to members of the Information Exchange; those within their organisations and/or constituencies (whether direct employees, consultants, contractors or outsource-staff working in the organisation) who have a NEED TO KNOW in order to take action.

GREEN Information can be shared with other organisations, Information Exchanges or individuals in the network security, information assurance or CNI community at large, but not published or posted on the web.

WHITE Information that is for public, unrestricted dissemination, publication, web-posting or broadcast. Any member of the Information Exchange may publish the information, subject to copyright.”

DNS Operations Analysis and Research Centre

(<https://www.dns-oarc.net/files/agreements/oarc-participation-201102.pdf>)

“6.3 Use of Information. All information submitted in connection with DNS-OARC shall be categorized as follows:

(a) Non-Confidential. Unless otherwise specified at the time of submission, all information shall be deemed non-confidential and may be disseminated to any parties, whether or not they are DNS-OARC Participants.

(b) Confidential – Analysis Only. If a Participant wishes certain information to be used for analytical purposes only, Participant shall clearly designate such information at the time of submission to DNS-OARC as “Confidential Information – Analysis Only.” The Board of Directors shall not disclose to DNS-OARC Participants such information in its raw state, but may review and analyze such information in accordance with DNS-OARC’s mission and purpose.

(c) Confidential Information – DNS-OARC Participants. If a Participant wishes certain information to be treated as confidential by the other DNS-OARC Participants, Participant shall clearly designate such information at the time of submission as confidential (“Confidential Information”). All such Confidential Information may be used by DNS-OARC and the other DNS-OARC Participants, provided that they (i) shall hold all Confidential Information in the highest confidence, (ii) shall not disclose, disseminate or publicize any Confidential Information to any person or entity, and (iii) shall take all action necessary or appropriate to hold all Confidential Information in the highest confidence. All Confidential Information shall remain confidential for a period of two (2) years after submission, unless a longer period is requested in writing by a Participant.

...”

UK National Hi-Tech Crime Unit (2003-2006)

(from an unofficial copy at

http://www.sourceuk.net/article/2/2476/confidentiality_charter_the_nhtcu_working_with_business.html)

“Reporting for Intelligence Purposes

The key aspects which will govern the way in which we work together on an ‘intelligence only’ basis are:

- First and foremost, all exchanges of information will be treated confidentially.
- Subsequently, when appropriate, we will agree with the company concerned, what information must remain confidential between us and what, if any, may be sanitised and disseminated for the benefit of industry in general.
- We will protect the source of information in accordance with our well-tried and proven source handling processes.
- We will distribute regular and timely business intelligence bulletins to inform industry of the latest trends and threats from hi-tech crime.

Reporting for Investigative Purposes

The key aspects which will govern the way we work together on an investigation are:

- We will actively support business risk assessment processes so that we understand the commercial impact and business sensitivities of any hi-tech attack.

- We will engage with the industry source and any affected third party before acting upon and disseminating information which has been provided.
- We will be sympathetic to the needs and priorities of business and, wherever possible, carry out enquiries in a way which will minimise disruption to the company concerned.

...”