Reacting to Cyberintrusions:
Technical, Legal and Ethical Issues

Richard E. Overill
Department of Computer Science
and
International Centre for Security Analysis
King’s College London
Strand
London
WC2R 2LS
U.K.

INTRODUCTION

The classical security paradigm of Protect, Detect, React has traditionally been applied to the field of information security with Firewalls taking on the role of protection while detection is handled by Intrusion Detection Systems (IDS). This admittedly simplistic picture leaves open two questions: who or what should react? and how?

While the role of reaction has traditionally been assumed by the system or network manager, it has become evident that an IDS which operates online and in real-time can also be configured to behave reactively or even proactively, thus taking over at least part of the manager’s role.

The potential behaviour of such an ‘active defence’ IDS raises at least three distinct kinds of issue for consideration [1]:

- technical possibilities — what behaviour is possible in practice;
- legal aspects — what behaviour falls within the appropriate legal framework;
- ethical considerations — what behaviour is acceptable in a particular social or business context.
TECHNICAL POSSIBILITIES

In principle there is a wide spectrum of potential responses which a ‘reactive’ IDS could make to a presumed intrusion. On a graduated scale ranging from benign to aggressive these responses might be categorised as follows:

- notify the operator / system manager / network manager by means of a console alarm / pager / email;
- send a warning email to the originator of the suspect process or connection;
- monitor and record suspect sessions or connections using system logs or raw network traffic data to provide forensic evidence or diagnostic material for future investigation;
- entrap the intruder into divulging identity information and other evidential material using a protective ‘sandbox’ or an enticing ‘honeypot’ as a decoy;
- discard a stream of suspicious network packets;
- terminate the suspect user process;
- disconnect the offending user connection;
- disable the affected user account;
- modify a router filter list to reject connection requests from the suspect IP source address;
- reconfigure the firewall to block requests for the particular IP service used by the suspected intruder;
- shut down and restart the affected machine;
- disconnect the affected node from the network;
- mount a denial-of-service reprisal attack using e.g. a HERF gun, worm, flood attack, etc. [2];
- launch a retaliatory malicious software strike using e.g. a virus, logic bomb, Trojan horse, etc. [3].

It is to be noted that three of the responses listed above are classified as ‘passive’ since the intruder should in principle remain unaware of them; the remainder are classified as ‘active’ responses, since the intruder will sooner or later become aware of having triggered them.

A ‘proactive’ IDS, on the other hand, might not wait to flag an intrusion but might instead take pre-emptive countermeasures:
• terminate all network connections or user processes which do not originate from *bona fide* users at approved sites [4], thus acting in the capacity of a firewall backup;

• mount a pre-emptive denial-of-service attack on ‘the enemy’;

• launch a pre-emptive malicious software strike on ‘the enemy’.

There are however potential problems with these ‘active defence’ strategies. In particular, unless the IDS detection thresholds are very finely tuned to minimise the occurrence of false alarms (or ‘false positives’) [5], a reactive IDS may be triggered by a naturally occurring false positive to disconnect an innocent user, or may be coerced by a strategically contrived false positive into unnecessarily shutting down a network connection.

These considerations were sharpened and focussed by the announcement in 1998 of Blitzkrieg [6] which was claimed to use self replicating (Worm-like) and self repairing (Core Wars) technologies. Two versions of this system have reportedly been developed: an aggressive military version is designed to wage cyberwarfare by launching malicious software attacks against intruders by attempting to damage or destroy information on their computers; a somewhat milder business version attempts to ward off denial-of-service and other common attacks where the intruder’s aim is to prevent the operation of a commercial service rather than to destroy data *per se*.

A further strand in the reactive countermeasures or ‘strikeback’ debate is the claim by a group of hackers known as the Electronic Disruption Theater (EDT) that their cyberattack in September 1998 on DefenseLink, the US DoD's primary public information Internet site, was cyberambushed by Pentagon officials. The Web browsers of anyone logging onto the EDT Website to participate in the denial-of-service attack, which was based on the Java applet FloodNet, were automatically shut down. The hackers claimed that this alleged form of offensive information warfare was illegal under the US Computer Fraud and Abuse Act, a charge which was rebutted by the DoD and the Pentagon [7]. However, it has been pointed out that a prime US military directive, “posse comitatus”, was breached; this directive forbids the US military from taking unilateral action within the US against US citizens [8].

**LEGAL ASPECTS**

These technical strategies also raise questions of legality. In the UK, the Computer Misuse Act 1990 [9] includes both a Basic Hacking offence and an Unauthorised Modification offence. Any attempt by an IDS to gain unauthorised access to an intruder’s computer would fall foul of the former offence. The launch of a malicious software strike or denial-of-service attack against an intruder’s system by an IDS is covered by the latter offence which carries a penalty
of up to 5 years imprisonment and/or an unlimited fine on conviction. Both
offences explicitly include trans-border modes of operation. These activities are
also illegal in the US under the Computer Fraud and Abuse Act 1986 and its
successor the National Information Infrastructure Protection Act 1996 [10].

The Council of Europe has produced a report highlighting the problems of
criminal procedural law connected with IT which contains recommendations
for legal principles on computer misuse [11]. Subsequently the Legal Advisory
Board of the European Commission commissioned a study on the legal aspects
of computer related crime in the information society [12] which produced specific
legal recommendations for the EU. Still more recently, the Council of Europe
has published the text of a draft Convention on cybercrime [13] with the aims
of harmonising national criminal legislation and facilitating co-operative inves-
tigations between member state authorities.

However, in the context of a military conflict between nation states, interna-
tional law, embodied in the 1945 UN Charter, does not resolve the ambiguities
that characterise information warfare activities. In particular, there is no real
clarification of the apparent conflict between the notion of sovereign nation
states and the reality of global digital networks [14]. Specifically, there is at
present no conclusive legal authority for what, if any, information warfare activi-
ties would constitute “armed attacks”, “aggression”, or “force” in international
law.

The 4th Geneva Convention of 1949, Convention Relative to the Protection
of Civilian Persons in Time of War, affords protection to individuals falling
under the jurisdiction of a belligerent, and includes a provision which outlaw
of collective punishments and reprisals. Whether this would include ‘collateral
damage’ to non-combatants and their vital infrastructures (such as hospitals,
power and water supplies) is open to question since the feature of intentionality
is absent in this case.

A similar comment applies to Article 17 of the 1950 European Convention on
Human Rights (as modified in 1998), which prohibits any state, group or person
from engaging in any activity or performing any act aimed at the destruction
or limitation of any rights and freedoms set forth in the Convention (except as
detailed) [15]). The Convention was enacted in the UK as the Human Rights
Act in October 2000 and it remains to be seen how its provisions are realised
in this setting.

In the IDS context, the use of ‘honeypots’ for enticing intruders in order to
determine their identities and monitor their techniques at close range raises an
interesting issue: it is at least possible that the use of a honeypot might be held
to constitute an incitement to commit a criminal act; as such it might render
the deployer, rather than the intruder, liable to prosecution.
ETHICAL CONSIDERATIONS

Equally important as the legal issues are the ethical implications of reactive countermeasures. 'Strikeback' has come to prominence [16] with the announcement of a report from WarRoom Research Inc. entitled Corporate America's Competitive Edge [17]. Released in January 1999, the report consolidates an 18-month study into cybersecurity and business intelligence issues. Of the 320 Fortune-500 companies surveyed, 30% claimed that they had installed software capable of launching counterattacks to security breaches. This appears to be a reaction to companies' previous reluctance to inform law enforcement agencies of security breaches for fear of unwanted public exposure with consequential damage to business confidence. While the companies in the WarRoom survey appeared to view strikeback as a right similar to the use of force in physical self-defence, they did not appear to understand its potential drawbacks.

Under certain circumstances an active IDS may retaliate against the wrong individual, or against someone who has made a genuine mistake or is harmlessly curious. In the latter scenario such behaviour might be considered unhelpful or even unethical. The former situation arises as a result of 'protocol spoofing' [18], where an attacker forges the IP source address of the network packets to make them appear to originate from an authorised user. If the contents of the packets are so constructed by the attacker that the traffic itself appears to contain an attack then the IDS will react [19]. This leads to the unfortunate possibility of dumping a legitimate user who has become the unwitting victim of an 'electronic framing' attack. A false positive flagged against a potential or actual commercial customer is likely to result in a consequential loss of goodwill and/or business. To quote David Curry of IBM “the last thing you want is to blow away a legitimate customer” [20].

Automated and concerted use of strategically contrived false positives and IP spoofing by an attacker may convince a reactive network-based IDS that an entire network subdomain is currently under attack and thereby subvert it into blocking legitimate network traffic, closing innocent network connections, or even shutting down the entire network subdomain. The resulting denial of service to employees and customers alike could have disastrous consequences for online transaction processing (OLTP) capability and cede a significant advantage to market competitors.

SUMMARY AND CONCLUSIONS

Verifying that a genuine intrusion incident has actually occurred can sometimes be extremely difficult. The cost of a verification failure may however be very high. Thus the risks associated with the occurrence of false positives and false negatives are both high.

We should not be seduced by the image of ICE, the Intrusion Countermea-
sures Electronics in William Gilson’s novel Neuromancer [21], into imagining that we can completely delegate the responsibility for reactive countermeasures to automated systems. Software-assisted reactivity under human control is the realistic limit of what can be achieved without dangerously compromising safety, legality or business ethics.

To conclude, the opening couplet from Robert Frost’s poem Fire and Ice [22]:

Some say the world will end in fire,
Some say in ice.

We have to ensure that we keep the fire outside our firewalls, but at the same time we must not allow the ICE to cause our downfall from within.

References


6


**Dr. Richard E. Overill, FBCS, FIMA** is a Senior Lecturer in Computer Science at King's College London (www.dcs.kcl.ac.uk/staff/richard/), and a member of the International Centre for Security Analysis (www.kcl.ac.uk/orgs/icsa/). His current research focuses on computer-related crime and information warfare, particularly Intrusion Detection Systems. This article is an expanded and updated version of a paper given at the 1st International Workshop on Recent Advances in Intrusion Detection (RAID'98), Louvain-la-Neuve, Belgium, 14th–15th September, 1998.