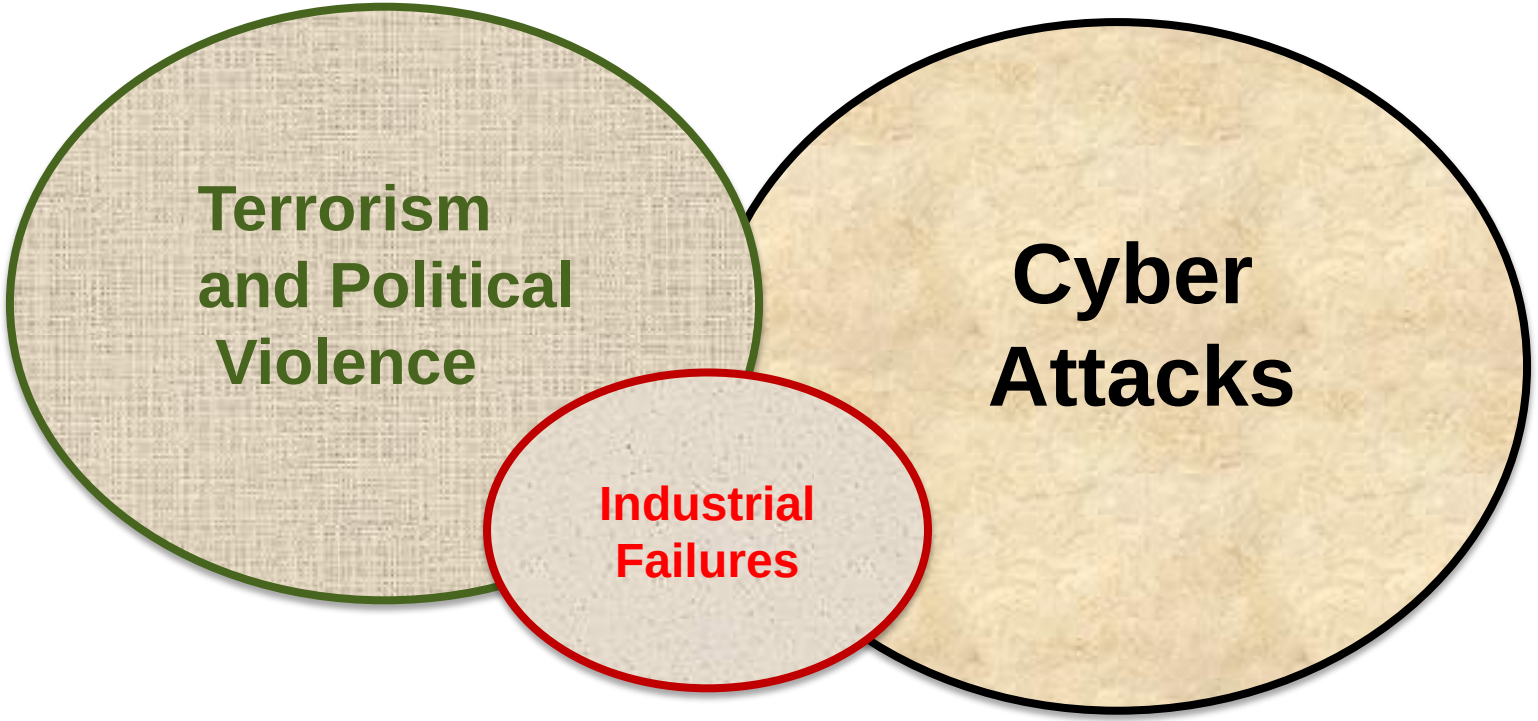
PROBABILISTIC MANMADE RISK MODELING

Dr. Gordon Woo

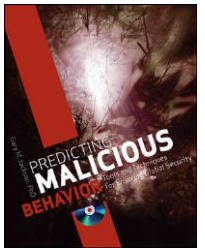
14 December 2016

Unified framework for modeling man-made catastrophe risk



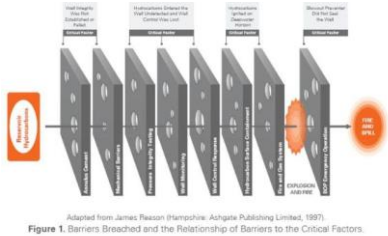
Principles for modeling man-made catastrophe risk

- The malicious behavior of terrorists, hackers and saboteurs, requires an adversarial **game theory** perspective.



- The universal prevalence of human error requires **probabilistic fault-tree analysis**.

- The common occurrence of near-misses requires **counterfactual disaster risk analysis**.

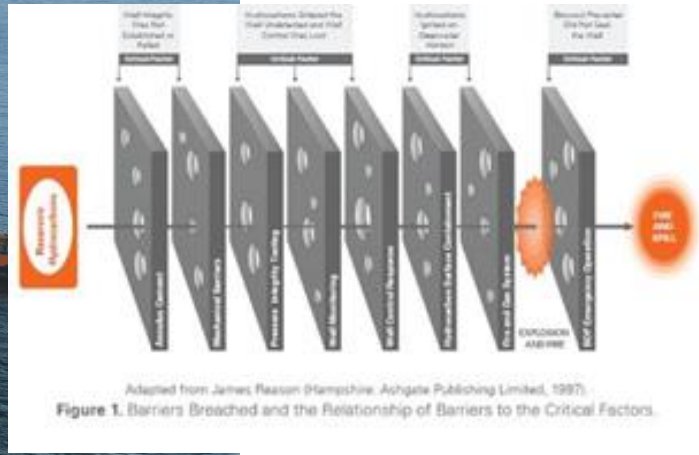


Adapted from James Reason (Hampshire: Ashgate Publishing Limited, 1997).

Figure 1. Barriers Breached and the Relationship of Barriers to the Critical Factors.

DEEPWATER HORIZON

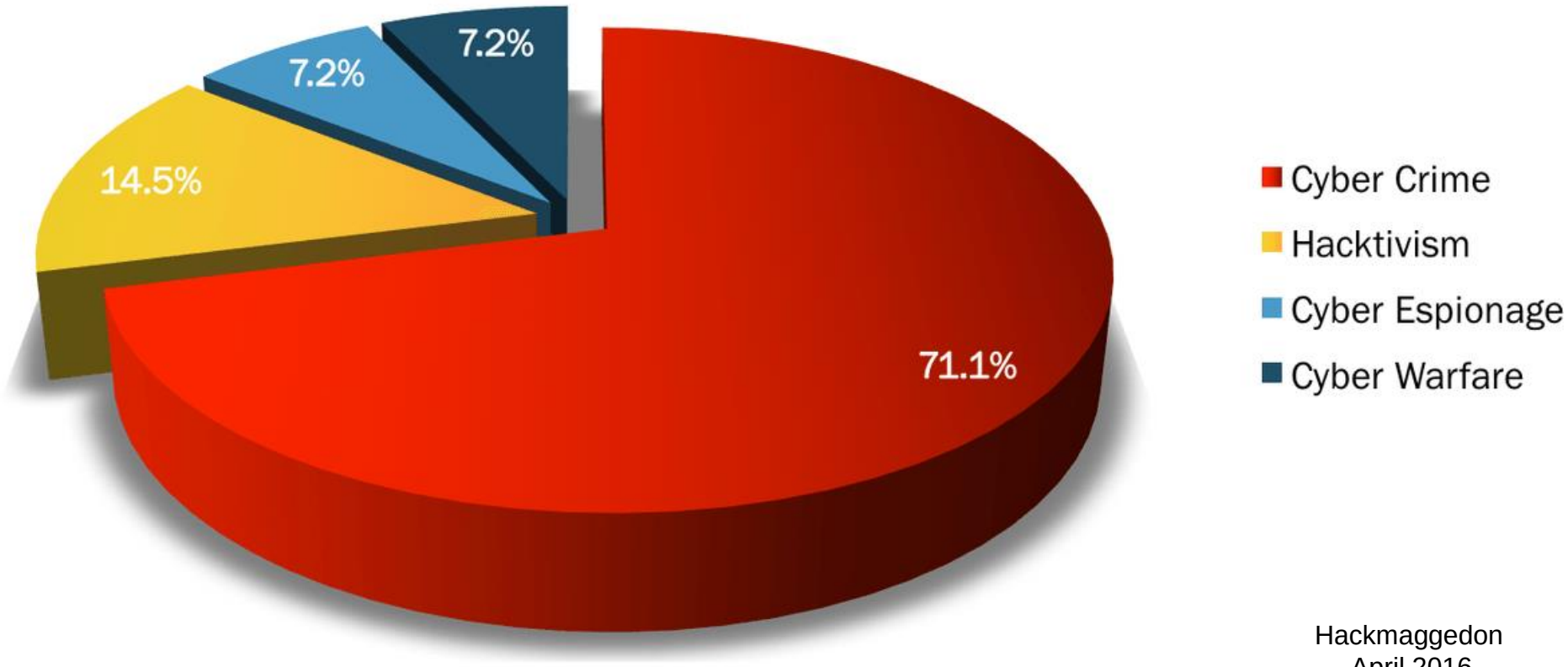
20 April 2010



Adapted from James Reason (Hampshire, Ashgate Publishing Limited, 1997).

Figure 1. Barriers Breached and the Relationship of Barriers to the Critical Factors

Motivations behind cyber attacks



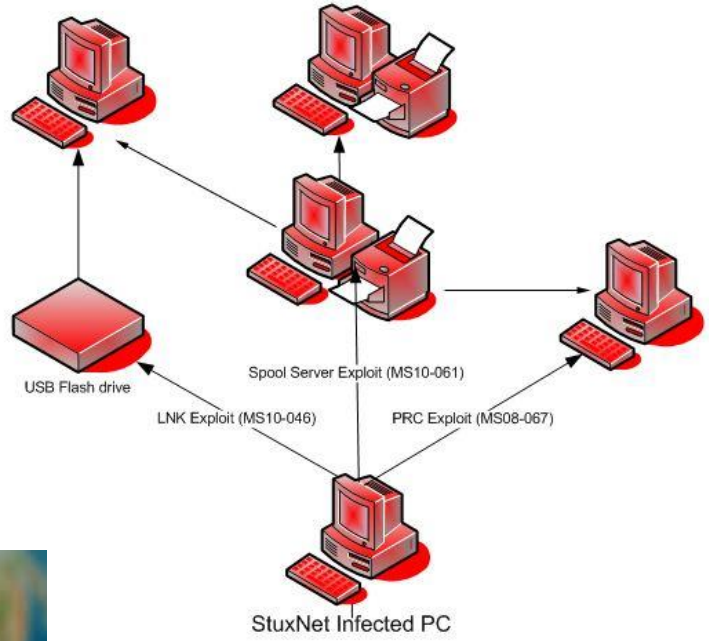
Hackmaggedon
April 2016

Hacktivism & activism

- Domestic terrorism is associated with political activism: social, economic, environmental etc.. Attacking involving computers is Hacktivism.
- Hacktivists mainly operate out of countries where they may be arrested, or from where they may be extradited.
- As with domestic terrorism, hacktivism is prone to ‘snitching’ by moles collaborating with law enforcement agencies.
- Hacktivist attacks involving physical violence overlap in targeting with domestic terrorism.

State sponsorship of cyber & terrorist attacks

- Cyber war is the fifth domain of warfare.
- State-sponsored cyber attacks can cause physical damage (e.g. StuxNet), as well as steal information.
- These are acts of state-sponsored terrorism – like the polonium poisoning of Litvinenko in London in 2006.



Terrorism: A Thinking Man's Game

'…This is a thinking man's game. Especially when one is as poor as the Popular Front is. It would be silly for us to even think of waging a regular war. We will continue our present strategy.
It's a smart one, you see…..'

George Habash (1970)

Following the Path of Least Resistance

‘The great principle is that,
in producing its effects,
Nature acts always according
to the simplest paths.

Pierre de Maupertuis, 1746

Threat Shifting

‘Now an army may be likened to water,

for just as water avoids heights, and hastens to the lowlands,

so an army avoids strength and strikes weakness.’

If two targets are equally attractive, terrorists will tend to attack that which has worse security.

Unlike with natural hazards, the likelihood of any building being targeted cannot be assigned independently of other buildings that might be targeted.

This also applies to human targets.

Alternative weapons of terrorist offence

'Terrorism is the language of being noticed.' Don DeLillo, writer

International name recognition in targeting

An important priority of terrorist targeting is international media coverage.

Terrorist attacks accordingly focus on targets in places with international name recognition, especially cities that are political, economic or tourist centres, New York, DC, London, Paris, Madrid, Mumbai, Oslo, Ottawa, Sydney etc..

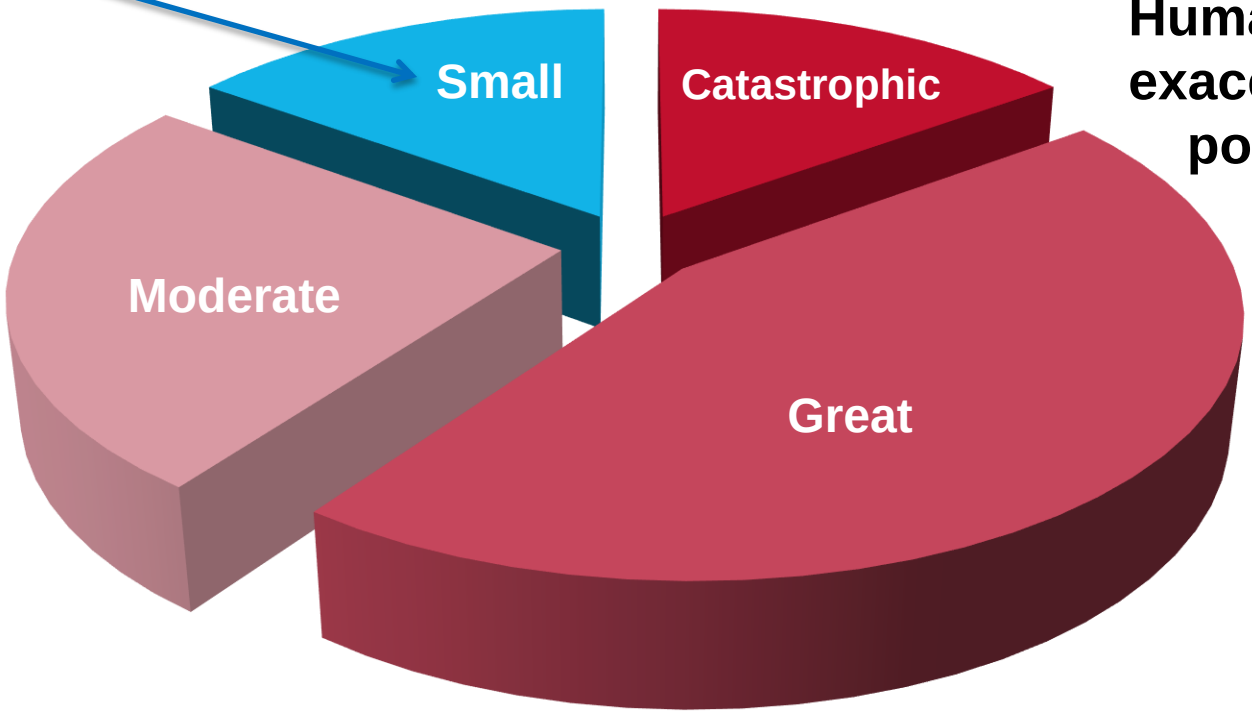
Soft targets in rural areas have minimal security - but also have poor media access.

'Terrorism without its horrified witnesses would be as pointless as a play without an audience.'

Mark Juergensmeyer,
Terror in the mind of God

Large scenario loss variability for man-made perils

Fortunate low historical loss outcome



Human factors exacerbate loss potential


16 January 2013: InAmenas, Algeria

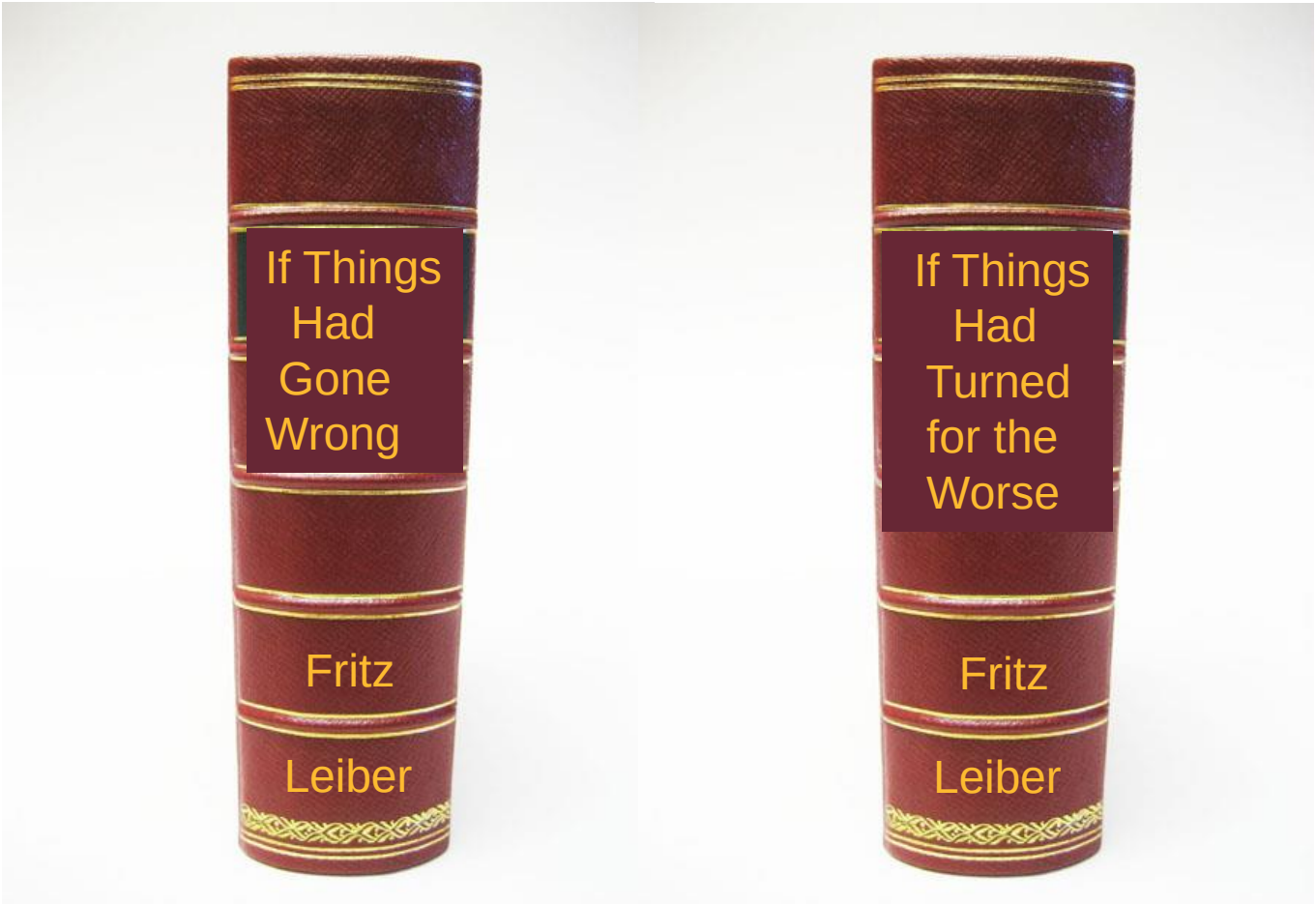
Aviation near-miss risk analysis:

AWE1549: 15 January 2009

Treatment of near-misses is crucial to assess aviation risk.

What if the pilot Sully had tried to return to La Guardia?

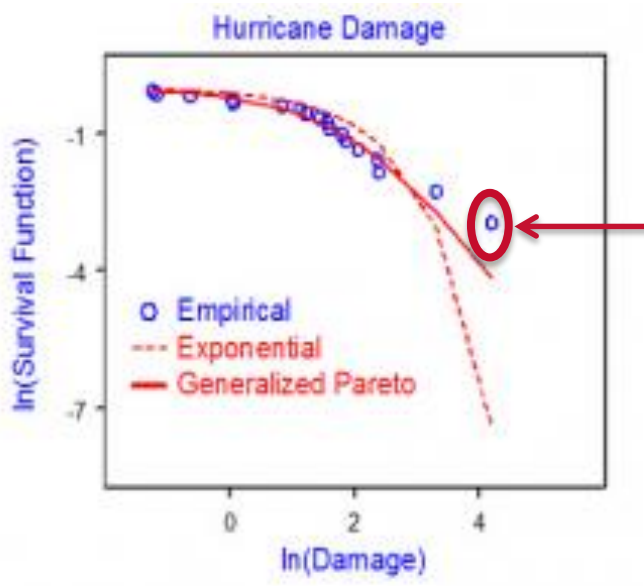
Studying downward counterfactuals



Counterfactual terrorism risk analysis

- In the western alliance, there are 15 years of terrorism loss experience since the era of catastrophe terrorism started on 9/11.
- This is about the same time span as the era of catastrophe cyber crime: the ‘I Love You’ virus caused about $10 billion computer loss in 2000.
- Due to effective counter-terrorism surveillance and law enforcement vigilance, there are few successful terrorist attacks – but numerous plots.
- Counterfactual analysis of these near-misses substantially extends the observation window of terrorism attacks. This allows the overall attack frequency to be estimated, as well as the attack mode probability, and targeting likelihood.

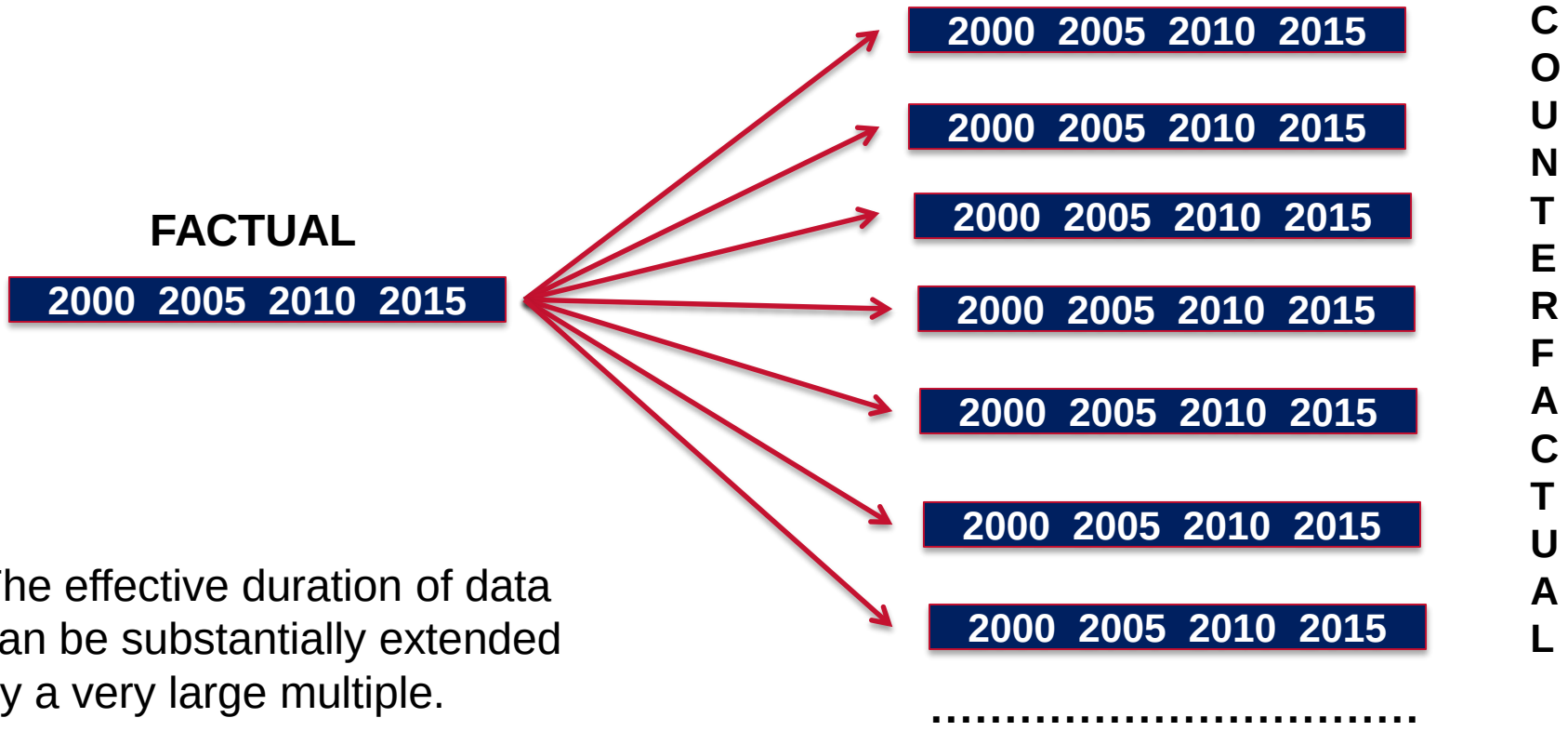
Statistical extreme value analysis



Statistical analysis of loss data is very sensitive to the largest losses in the historical database.

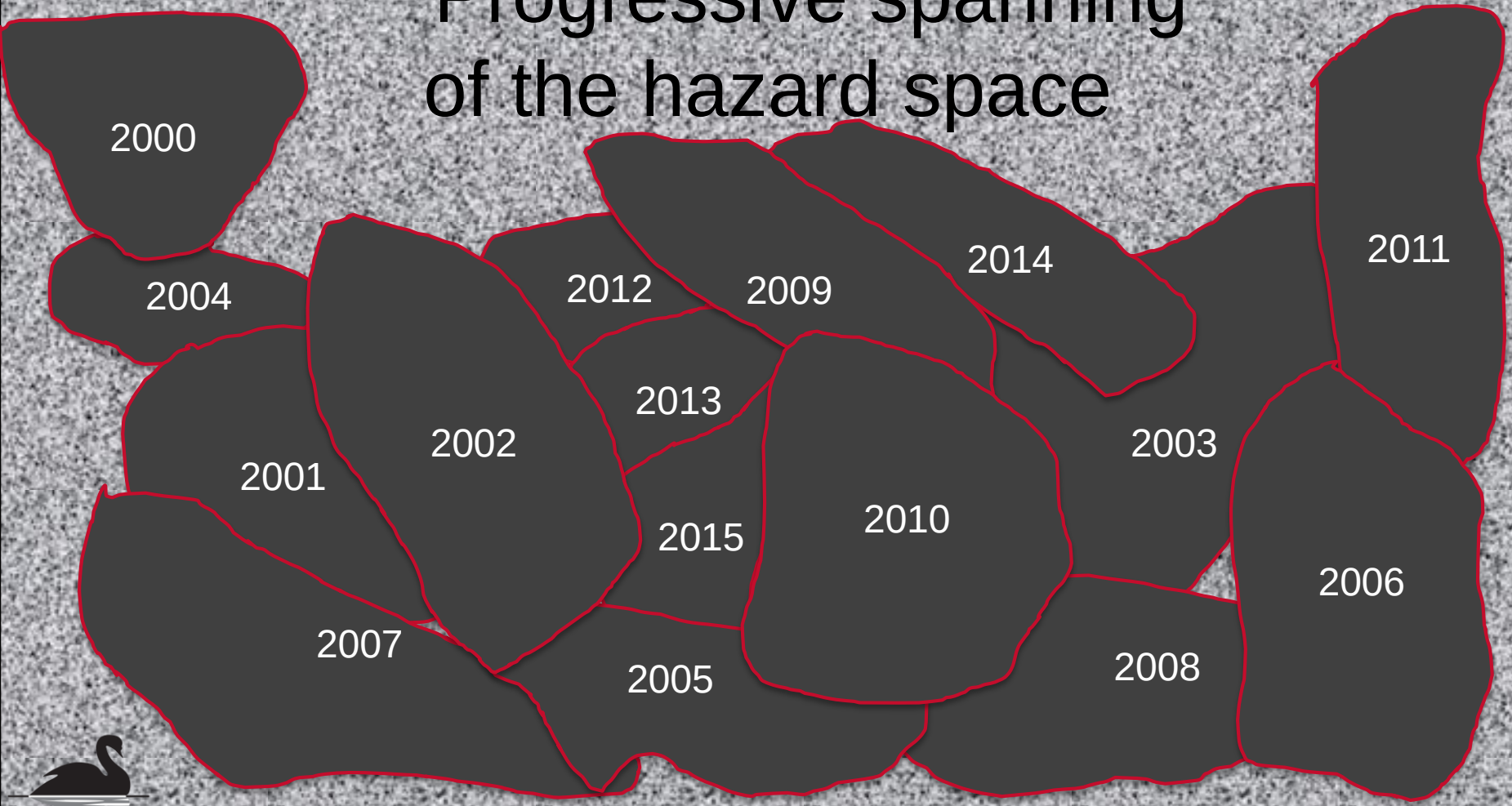
The duration of the historical database may be comparatively short, and the tail of the loss curve may include few major loss events.

Multiple alternative realizations of the past



The effective duration of data can be substantially extended by a very large multiple.

Progressive spanning of the hazard space



Counterfactual cyber risk analysis

- The space of alternative stochastic realizations of the hundreds of large cyber attacks over the past 15 years should span a high proportion of the cyber threat space.
- The potential losses for a particular attack are subject to very large variability due to the global scaling characteristic of cyber attacks, changing vulnerability profiles, and the large variability in footprint sizes.
- An assessment of the counterfactual loss distribution for the most significant historical cyber attacks provides insight into the probability of extreme losses.

SWIFT banking system heist

Malware was planted on the computer systems of the Bangladesh Bank.
A series of cloned transactions were executed via the New York Fed
to accounts in Sri Lanka and the Philippines.

On February 4, 2016, **$81m** was stolen from the Bangladesh Bank
and was laundered in the Philippines casino industry.
Another $20 million was recovered due to a Sri Lanka bank query:
'*Foundation*' being mispelt as '*Fandation*'.

A total of **$951m** might have been lost, except that 30 of the 35 orders worth $850m were blocked by the New York Fed as being suspicious.

This was a fluke: the street address of the recipient Philippine National Bank included the word 'JUPITER', which was the name of an oil tanker on an Iranian sanctions blacklist.

Home Depot Point-of-Sale Data Loss

- At Home Depot, around 56 million debit and credit card details were leaked in a breach that lasted from April to September 2014.
- The cyber thieves broke in using credentials stolen from a third-party vendor. These credentials did not provide direct access to POS devices. *A zero day vulnerability in Windows was needed*, which gave elevated rights to navigate the Home Depot network.
- The intruders targeted 7,500 self-checkout lanes because these were clearly referenced as payment terminals.
- Another 70,000 regular terminals which were identified simply by a number might also have been attacked. Furthermore, the data exfiltration period might have been extended a few months longer.

Home Depot Point-of-Sale Data Loss

- Denote the maximum possible data loss as L(max).
- Normalize exfiltration loss to L(max).
- The actual loss was about 10%.
- What is the chance that the loss might have been higher, e.g. 25%, 50% etc.?
- There are strategic game-theoretic constraints on the size of loss, given the risk of intrusion discovery, and the black market for selling data.
- How can this loss distribution be estimated?

Cloud service provider failure

A major crash at Amazon Web Services occurred on April 21st, 2011. In addition to taking down the sites of dozens of high-profile companies for hours - and even days, the crash permanently destroyed some data. A detailed diagnostic report was produced a week later.

'At 12:47 AM PDT on April 21st, 2011, a network change was performed as part of our normal AWS scaling activities in a single Availability Zone in the US East Region. During the change, a traffic shift was executed incorrectly leading to a re-mirroring storm.'

The direct cause of the crash was a human error - but it might have been a malicious act, with worse consequences. The task of recovering from this major storm, with a minimum loss of data, was a highly complex challenge. The recovery time distribution would have had a long tail.

'The Amazon interruption was the computing equivalent of an airplane crash.' New York Times

IoT DDOS: October 2016

The source code that powered the “Internet of Things” botnet responsible for launching the DDoS attack on Dyn has been publicly released.

The Internet may be flooded with attacks from many new botnets powered by insecure routers, IP cameras, digital video recorders and other easily hackable devices.

The malware, **Mirai**, spreads to vulnerable devices by continuously scanning the Internet for IoT systems protected by factory default or hard-coded usernames and passwords.

A seller is offering an IoT botnet capable of generating 1 terabyte/sec. [$7,500 buys 100,000 computer botnet].

Heartbleed: April 2014

Heartbleed was a critical bug (CVE-2014-0160) in the popular OpenSSL cryptographic software library.

The heartbleed bug could result in disclosure of the contents of a server's memory, where the most sensitive data is stored; including usernames, passwords, and credit card numbers.

Fortunately, this bug was first reported to the OpenSSL team by Google Security. (It was independently discovered at the security company Codenomicon).

However, if this bug had been discovered first by a malevolent hacker, there might have been a severe data exfiltration loss to some otherwise secure major organizations.

Shadowbrokers NSA cyber weapons

According to NSA, a former agency operative carelessly left the hacking tool files available on a remote computer, where Russian hackers found them. NSA did not detect use of the tools by China and Russia, and NSA did not feel obligated to warn the US manufacturers.

The New York Times reported that, on August 27, 2016, the FBI arrested Harold Martin, who worked at NSA. In his home were found many terabytes of highly classified information from the 1990s until 2014.

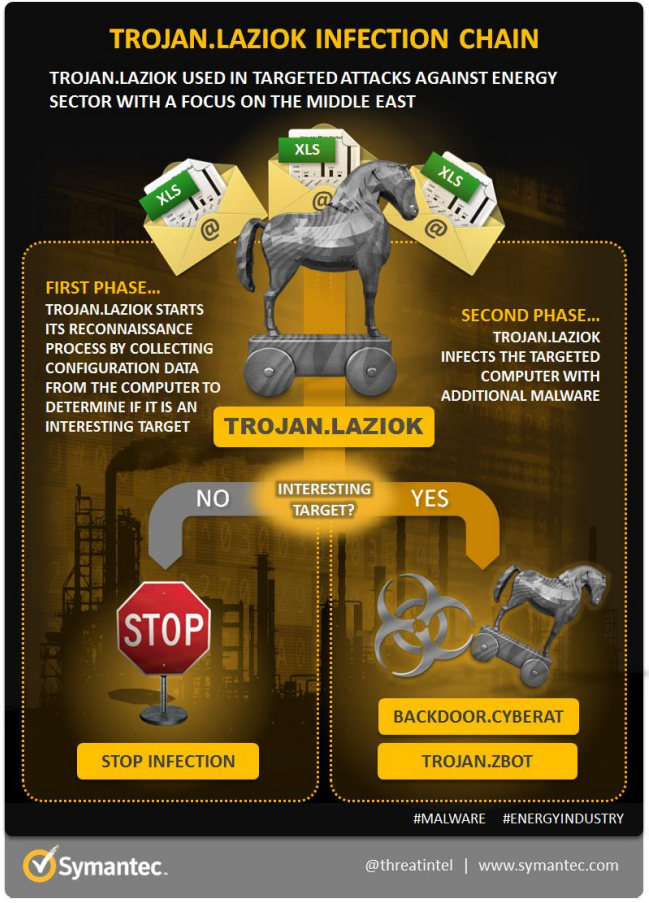
Regardless of the cause, whether negligence or theft, either contingency shows the serious possibility of a release of weapons from the NSA (Equation Group) cyber arsenal.

From espionage to sabotage

In early 2015, a Trojan (Laziok) targeted firms in the energy industry, infiltrating systems to gather information about operations.

The stolen data enabled the attacker to make decisions about how to proceed with the attack, such as installing additional Trojans and backdoors.

Once espionage malware has entered infrastructure, it can send more advanced malware to execute damaging attacks on infrastructure.



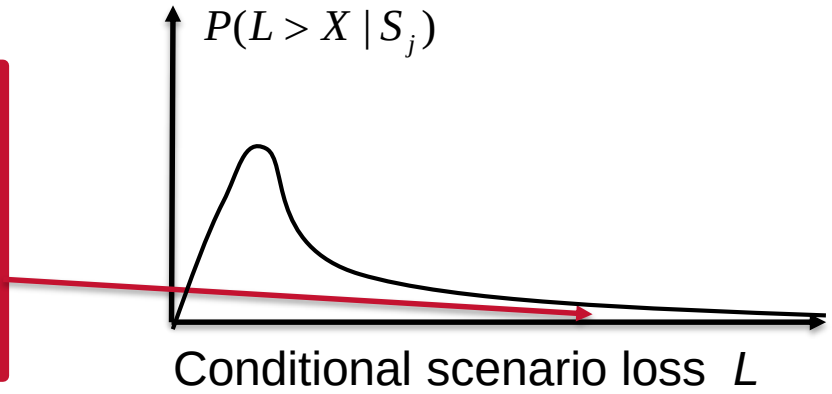
Mathematical framework for tail risk modeling

Define a set of cyber attack scenarios, indexed by category of attacker (e.g. hacktivist, Russian state), and attack mode (e.g. DDOS, exfiltration).

The annual frequency of the j'th attack scenario S_j is written as: $f(S_j)$

$$P(L > X) = \sum_j f(S_j) * P(L > X \mid S_j)$$

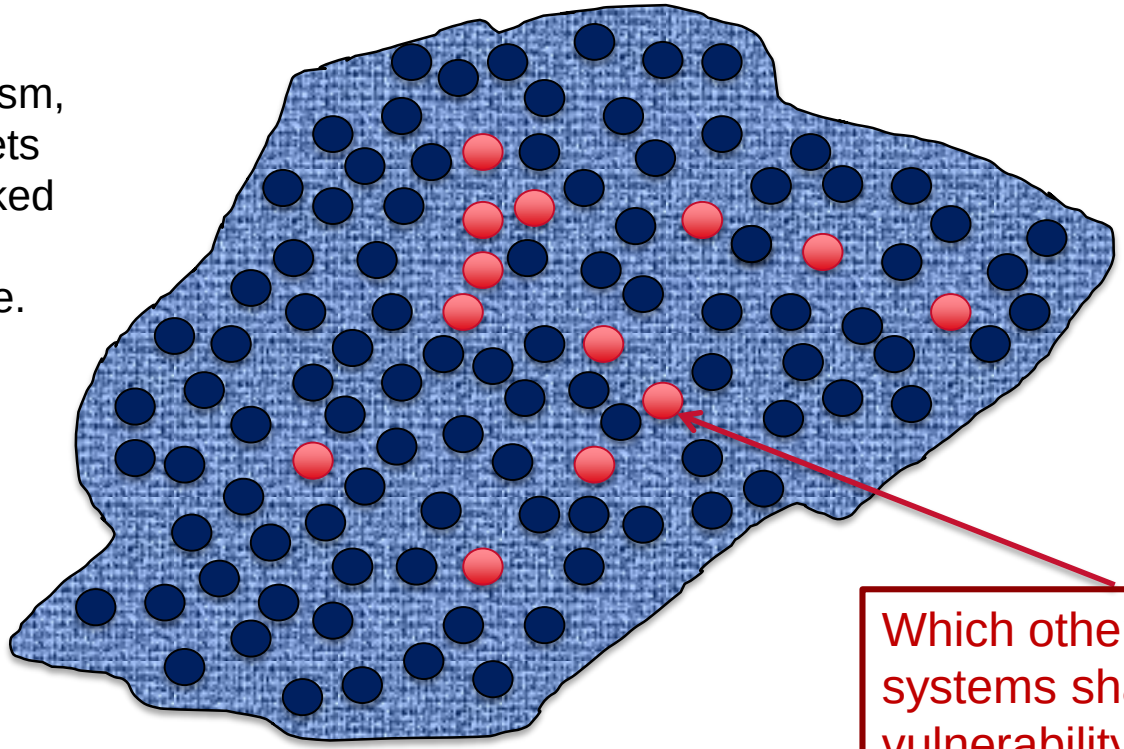
An extreme loss can arise from an attack scenario which has happened in the past 15 years, and where the loss is in the far tail of the broad conditional loss distribution.



Scaling of footprints in IP space

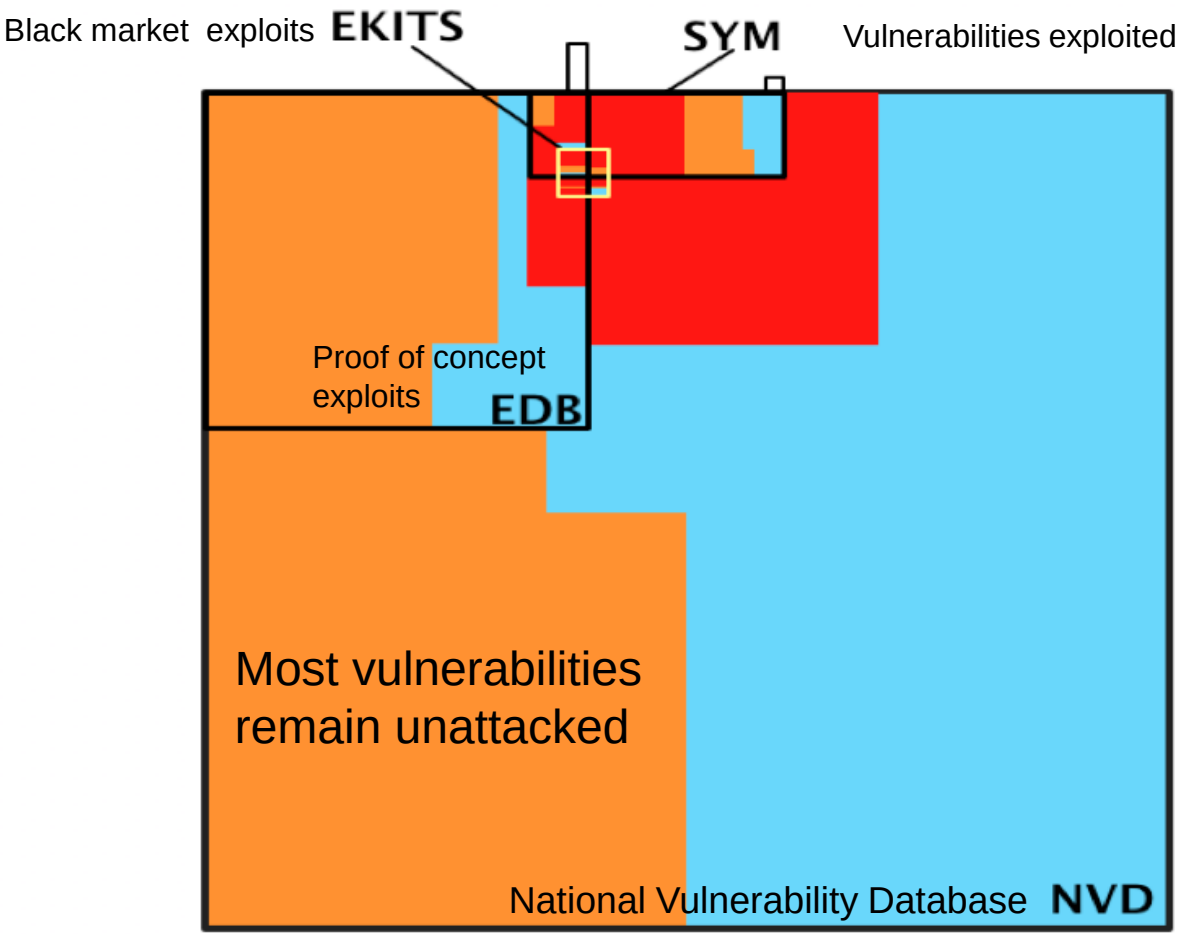
In contrast with conventional terrorism, the number of targets that might be attacked can scale by many orders of magnitude.

This lengthens considerably the tail of the conditional cyber loss distribution.



Which other computer systems share the same vulnerability, and how many become infected?

Attacker: Russian criminals
Attack mode: Data exfiltration



Coloured areas are proportional to data size.

High score vulnerabilities with (CVSS ≥ 9) are in red;

Medium score vulnerabilities with (6 ≤≤ CVSS < 9) are in orange;

low score vulnerabilities with (CVSS < 6) are in cyan.

(Allodi & Massacci, 2013)

LOW CVSS<6 | 6≤MEDIUM CVSS<9 | HIGH CVSS ≥ 9

Common **V**ulnerability **S**coring **S**ystem

Blackhole 1.2.0: a very popular exploit kit

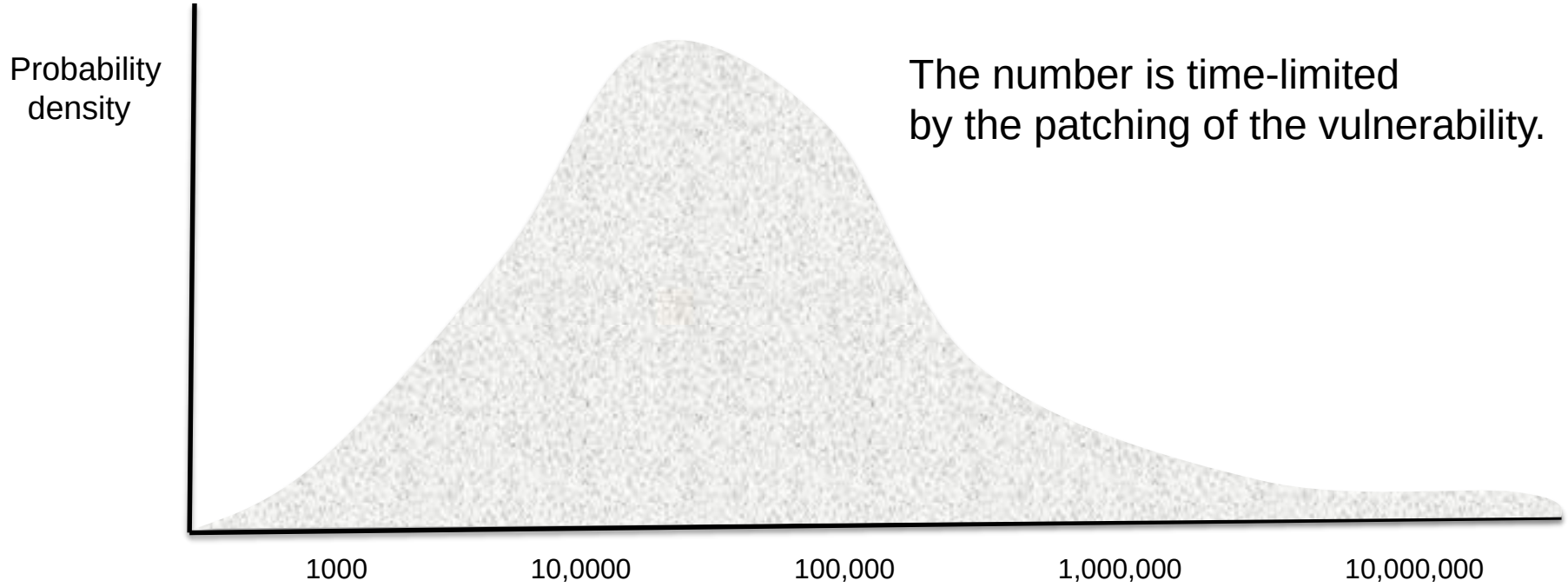
In September 2011, there was a peak in the relative risk from EKITS coming from a very popular Russian exploit kit: Blackhole 1.2.0. This EKIT exploited the CVE-2012-5076 vulnerability, which allows an attacker to download and install files onto a computer with a vulnerable version of Java.

If a website containing the malicious code is visited, while using a vulnerable version of Java, the exploit code is loaded. It then attempts to download and execute files from a remote host/URL. The files that are downloaded and executed could include additional malware.

It is known that about 40% of attacks were from black market EKITS in September 2011.
A long-tailed probability distribution for the number of infected computers from Blackhole 1.2.0 can be obtained by estimating the following parameters, with associated uncertainties:

- The number of computers having the vulnerable version of Java
- The proportion that would visit a website with malicious code, or download infected email files
- The prevalence of inadequate security measures to deal with the malware.

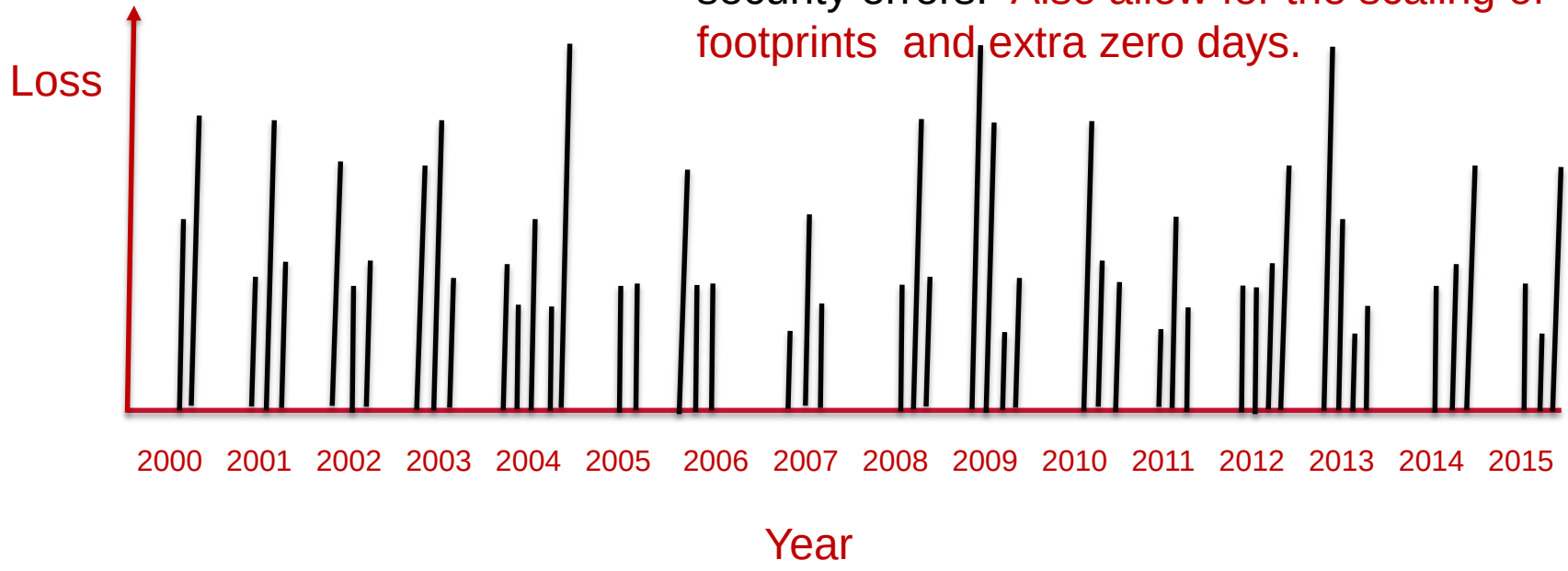
Scalability probability distribution for an infection



Stochastic simulation of past major events

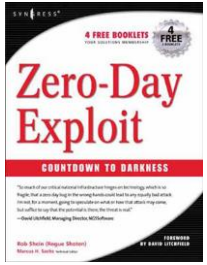
Attacker: Russian criminals
Attack mode: Data exfiltration

Sample the intrinsic uncertainty in target loss, accounting for intrusion detection and other defensive counter-measures, as well as human security errors. Also allow for the scaling of footprints and extra zero days.



Recipe for re-sampling major historical events

Repeat adding zero days



This provides the basis for a viable compact stochastic cyber risk model, which can be supplemented with some extra extreme events.

Each major historical event spawns an ensemble of significant future events.

Probabilistic cyber risk model steps

- Based on counterfactual analysis, develop a statistical model for the uncertainty in cyber loss at a specific target, for different categories of attacker and modes of cyber attack.
- Obtain statistics on malware infectivity: waterholing, phishing etc..
- Acquire data on historical cyber attack footprints. Acquire IT information on corporate computer systems and security.
- Adapt an economic game-theory model for the discovery, hoarding and deployment of zero days by different organizations, and for alternative attack modes.
- Construct a basic set of stochastic simulations for probabilistic cyber risk analysis.