Information and Information Security

JAAK HENNO, Tallinn University of Technology

Tremendous growth of Information Security issues poses a problem - is most of mankind turning criminals or is something wrong with rules? Are we always trying to secure valuable Information or just data and do the rules correspond to generally accepted behaviour?

In order to understand better issues connected with Information Security we have first establish concise meaning of terms 'Information' and 'Information Security'. Information cannot be considered separately from an Information Processing System (IPS); a message is Information only for some IPS and is used by the IPS for achieving its goals.

In the paper is considered general model of IPSs and their goals. To get some insight of attitudes of users, the mileu of Massively Multiplayer On-line Games (MMOG) is considered. A 'Security Incident' may be just a curiosity; cheating in a game is not considered serious offence by fellow players.

Categories and Subject Descriptors: **H.1.0 [General]:** Information/Entropy—*Information/Processing*; **H.1.1 [Systems and Information Theory]:** Natural Information Processing Systems—*Unified view/methodology*; **H.1.2 [Models and Principles].**

General Terms: Information, Entropy, Information Processing Systems, Information Security

Additional Key Words and Phrases: data, models, games, MMORG

1. INTRODUCTION

Insecurity, threats to information, need to defend information from unauthorized access, use, disclosure, disruption, modification, perusal, inspection, recording or destruction are nearly always deliberate. Accidents, which in 90ties were responsible for most Information Security issues - a plane flows into an office building destroying all computers in it, employee mistakes ("Format C:"), fire, flood, earthquake, lightning, shooting or otherwise destroying a computer in fit of anger ("You do not have right to access folder 'My Documents'!), issues with ISP or WAN etc account currently only for tiny part of Information Security problems. According to recent report from Panda Labs [PandaSecurity 2014], in 2013 appeared about 82,000 new malware threats per day and in the whole year - 30 million new malware threats, Kaspersky Lab is detecting 315,000 new malicious files every day [Kaspersky 2014] most of them (>70%) - trojans, especially designed for stealing/damaging Information.

Current computerized Information Processing systems are vulnerable, rigid and not adaptive, since the main focus in their development is on computer technology and communication protocols. System environments, culture of system users, reasons for attacks, culture and operating modes of attackers are considered less.

Threats to Information are not specific only to computers and computer networks, they are present in nearly any Information Processing System (IPS) - social and business organizations, governments, all kinds of living systems down to simplest ones - cells and bacteria. In order understand mechanisms which provoke misuse of information and design and create adaptive information security systems, which can adequately respond to constantly changing dynamic environment and threats and can secure functioning of IPS under attacks and threats the IPS should be considered on more general level. But we should also re-consider rules and practices – do they agree with general understanding of right and wrong?

Author's address: J. Henno, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn; email: jaakatcc.ttu.ee

In: Z. Budimac, T. Galinac Grbac (eds.): Proceedings of the 3rd Workshop on Software Quality, Analysis, Monitoring, Improvement, and Applications (SQAMIA), Lovran, Croatia, 19.-22.9.2014, published at http://ceur-ws.org

2. INFORMATION FLOOD

Information regulates all processes and information processing is the central characteristics of all Universe. Advances of Biology, Economy, Information Technology, Social studies etc have introduced and supported the understanding, that we are living among Information Processing Systems Ko Seth Lloyd from MIT calculated in 2006 [Lloyd 2006], that the visible universe has so far computed about 10^{122} operations on 10^{92} bits.

Even a simple unicellular organism is a complex and purposefully organized algorithm, but the most complex IPS are humans. Man is the most complex information-processing system on Earth. By some estimates [Spiro 1980] the human brain can store 2.5 petabytes (ca million gigabytes); in the human body are in every second processed 3.4×10^{19} bits.

Our individual Information Processing capabilities are tremendously amplified by our Society and our greatest invention - Internet; so much so, that we clearly are not any more able to manage the process. The amount of digital data created in World grows in two years (at least) twice and in 2020 will be 40000 exabytes (exabyte = 1018 bytes) - over 300 times more than it was in 2005 [EMC 2014]. World population in 2020 is estimated to be 7.72 billion, thus the memory capacity of the whole mankind can in 2020 store only 0.00004% of the whole digital data generated in our digital universe.

We believe, that we are rational beings and everything we do has a purpose. What is the purpose of this flood of data ? Some people call this 'Information Age', 'Age of Big Data', but what is the purpose of amounts of data, what we can not manage? Software is eating the world [Andreessen 2011], who will be in control in 2020 - we or our programs? How much of this data is Information, what has for us some use? How much of this should be protected using the Digital Rights Management (DRM) methods and is this protection in interests of whole Mankind?

3. WHAT IS INFORMATION?

We all agree that we now live in "Information Age", we have become an "Information Society" and that information is the main source of value in the global economy. Several scientists have proposed [Stonier 1990, Hefner 1992, Gitt 1994], that that information is a part of the physical universe the same way as matter and energy. Information present in a system is the measure of organization of the system.

But in spite of vast number of papers on Information, Information systems, Information Security etc the meaning of the word "Information" remains abstract and underdefined. The situation has not become essentially better from the famous utterance from father of Cybernetics Norbert Wiener: "Information is information, not matter or energy" [Wiener 1948, p. 132]. In spite of proliferation of information systems, there is still no generally agreed answers to the questions – What is Information? Has Information natural properties and if so, then what are these properties?

The word 'information' is derived from the Latin word 'informare' - "give form to", i.e. information is always represented with some form, structure, pattern. But 'representation' of a concept (e.g. in human mind) is not the concept itself, representations are the result of grounding our sensory perceptions, discovering patterns in perceptions.

The Oxford Dictionary explains 'information' as 'facts provided or learned about something or someone', i.e. information is communicated to receaver.

It is often claimed that Information Theory was created by Claude Shannon, but the main topic of Shannon's research was Communication. Shannon always considered communication and did not speak about information; he always used notions 'communication, communications channel'. In his ground-breaking paper "A Mathematical Theory of Communication" [Shannon 1948] he explicitly states: "fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point", but "semantic aspects of communication are irrelevant to the engineering problem". Shannon considered messages and communication without any assumptions about their meaning, but information is based on meaning. A message without meaning does not contain information for receiver.

Shannon's paper started the 'Information Theory' boom in many fields- linguistics, biology, physics, so that he was forced to warn: "Workers in other fields should realize that the basic results of the subject [communication channels] are aimed in a very specific direction, a direction that is not necessarily relevant to such fields as psychology, economics, and other social sciences".

Meaning of communicated signals, how these signals are used in receiving IPS both in mechanical and biological systems, i.e. how signals influence, control these systems this was considered by Norbert Wiener considered in his ground-breaking opus "Cybernetics or Control and Communication in the Animal and the Machine" [Wiener 1948]. Wiener considered signals as information. Signals, data become for receiver information if they make sense for receiver and receiver can use them for achieving its goals. A data item is Information for an IPS if it helps the IPS in achieving its goals, i.e if it changes the future behaviour of the IPS, acts for the IPS like a program.

4. INFORMATION PROCESSING SYSTEMS

All Information Processing systems (IPS) - living systems, businesses, social organizations, governments, languages, computer programs, etc are finite and have goals, their purpose is to perform some actions aimed to fulfil their goals and they survive by constantly processing information about threats and opportunities in the world around them.

Goals of living systems are metabolism (getting energy needed for their functioning), growth and reproduction, secureing resources/territory and collaboration, goals of businesses - acquiring resources (money) by production of some goods and/or services, growth and creating daughter businesses, goals of languages - improve communication of language users, goals of governments - guarantee and improve well-being of citizens etc. Actions of IPS systems are based on information they receive from their environment; these actions allow them survive and develop, become more complex, reduce their inner entropy. Thus all IPS have (at least) these goals:

1. *metabolism - acquiring energy needed of its actions*
2. *growth/reproduction*
3. *secureing their existance, resources/territory from external threats*
4. *collaboration with other similar IPS - when this helps in achieveing their goals.*

Problems with Security arise from the third goal.

4.1 Formal Level

Actions for achieving of goals can be performed only if some conditions are true. Conditions can be expressed with predicates, thus we can consider goals as logical predicates.

Signals, data becomes information for an IPS only when they can be used for proving true their goals. A goal for a living system can be e.g. an eatable object:

$$a \wedge eatable(a)$$

For processing (proving a goal predicate true) data should be stored by IPS. Since all IPS are finite, they tend not to store useless data or if something is stored, but could not be used for a long time, then it is forgotten. According to Landauer's erasure principle [Landauer 1961] forgetting, erasing memory always increases thermodynamic entropy in the environment [Plenio 2001].

If the goal is predicate *vitamins* and IPS has already some information, e.g.

$$vitamins \Leftrightarrow \exists x(vitamins(\mathrm{x}))$$

$$vitamins(x) \Leftrightarrow (apple(\mathrm{x}) \vee pear(\mathrm{x})) \neg chips(\mathrm{x})$$

and IPS gets from its environment grounded information

$$a \wedge apple(a)$$

then it can deduce the goal $vitamins(a)$ and perform the action - $eat(a)$.

Shannon 's formula

$$H(x) = -K\sum_i p(x_i)\ln(p(x_i))$$

for calculating entropy of the used above function

$$f(x,y,z) = (x \vee y) \barwedge z$$

gives

$$H(f) = -(\frac{3}{8}\ln\frac{3}{8} + \frac{5}{8}\ln\frac{5}{8}) = 0.954$$

When IPS performs the above deduction and thereafter clears the used memory (its finite!), its entropy increases by this amount.

5. EXAMPLE TO CONSIDER - MMOGS

Problems with Security for environment arise from the third goal of IPS - securing their existence, resources/territory from external threats. IPS should keep balance with their environment for sustainable existence, but humans are often driven by greed and often overact, trying to secure more than inevitable, cheat and commit fraud. Every Security incident means, that somebody is violating the accepted rules and practices. The tremendous growth of security accidents poses a problem: wheather the whole Mankind is becoming criminals or is something wrong with rules, the rules do not agree with what is considered normal.

Reasons of our behaviour lay deep in human culture, history and psychology and are difficult to analyse, but for illustrative 'sandbox example' could be used computer games, especially Massively Multiplayer Online Games (MMOG).

The online gaming industry has in the past decades grown rapidly. The virtual economies created within MMOGs often blur the lines between real and virtual worlds so that several companies producing MMOGs have hired an economists to regulate the virtual economy in their games [Plumer 2012].

As a result, the Security problems in MMOGs are becoming increasingly critical. Cheating, virtual frauds, and other security attacks are widespread in the virtual worlds of MMOGs. Gamers have developed illegal methods to obtain virtual wealth and convert it into real wealth in the real world using Real Money Trading, which allows to buy using real money game values (tropies, achievements).

MMOGs are played in massively distributed systems spreaded over many countries with thousands of client processes interweaving on common servers in real time. This greates many security issues and is the reason for the lack of regulations in this industry - it is difficult to impose legal laws in a world that is virtual, "not real", involving several countries.

As a result has appeared a million-dollar business of developing cheating tools, bots for (half) automated gaming, cheats, trainers and walkthroughs and web-sites, where one could get for monthly fee super-human abilities in some game, e.g. see other players through walls [Maiberg 2014].

It is difficult to tell exact percentage of cheating players. In a reacent study involving one on-line racing game [Blackburn et al 2014] was collected information about gamers on the Steam Community global gaming network: Steam has an anti-cheating system which marks cheaters public profile; from more than 12 million analyzed gamers over 700 thousand (ca 6 %) had their profiles flagged as cheaters.

Cheating was earlier considered as a despicable behaviour, but currently the attitude seems to be changing - the study discovered, that "cheaters are well embedded in the social and interaction networks: their network position is largely indistinguishable from that of fair players". And seeing fellow players succeed through cheating creates the idea that it might be worth the risk also to try.

Cheating is often difficult to distinguish from healthy curiosity. Consider a simple on-line Pong game, where player has to move his paddle in order to bounce a ball to target of another player. If player can place its paddle in correct position (ball bounces from paddle), he/she earns some points; if ball misses the paddle, it also bounces back, but player does not earn anything.

Fig. 1. Cosmic Pong - player has to move his paddle to bounce the ball back and earn points; if ball misses paddle, it bounces back from ground and player does not earn anything.

We did such a game in game-programming class. Everything went well until one student started to earn rapidly tremendous scores. It turned out, that he has decreased the y-coordinate of his paddle, so that ball often started to bounce back from the bottom of the paddle creating several bounces, which quickly increased his score.

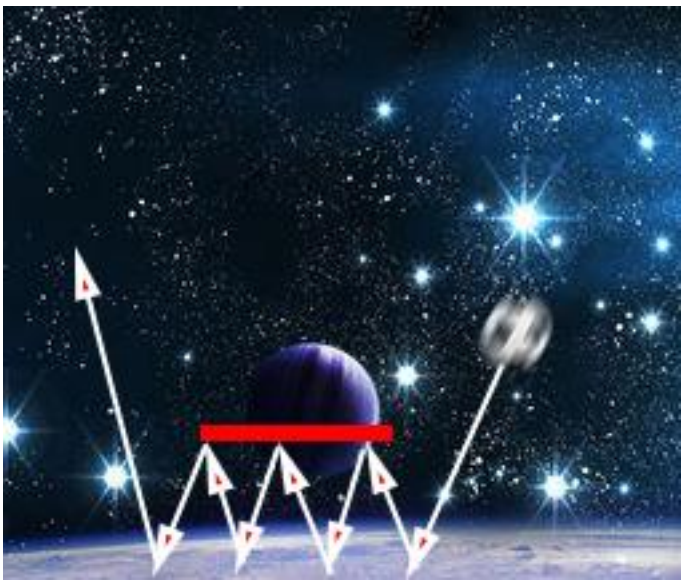

Fig. 2. Student's modification - paddle's y-coordinate is decreased, so that ball starts bouncing from the bottom of the paddle and creates every time new points.

What should teacher do - prize for ingenuity or punish for cheating, his modification give him clear advantage? After discussion we changed the game: everyone can modify (some) constants, but the modification will be at once distributed to all computers, so that all players are all the time in equal conditions.

REFERENCES

Andreessen, M. 2011. Why Software Is Eating The World.Brandtzæg, P.B. 2013. Big Data, for better or worse: 90% of world's data generated over last two years. Sciencedaily. (2013).

Burgin, M. 2010. Theory of Information 2010. World Scientific. ISBN-978-981-283-548-2 (2010), 283–548.

Duncan, T.L. and Semura, J.S. 2007. Information Loss as a Foundational Principle for the Second Law of Thermodynamics. Foundations of Physics. 37, (2007), 1767–1773.

EMC. EMC Digital Universe Study (2014), http://www.emc.com/leadership/digital-universe/index.htm

Galbraith, J.R. 1973. Designing Complex Organizations. Addison-Wesley.

Gánti, T. 2003. The Principles of Life. Oxford University Press.

Gantz, J. and Reinsel, D. 2012. The Digital Universe in 2020. JDC.

Gitt, W. 1994. Am Anfang war die Information. Am Anfang war die Information.

Hefner, K. 1992. Evolution of Information Processing Systems. Springer.

J. Blackburn, A.I. N. Kourtellis, J. Skorvetz, M. Ripenau 2014. Cheating in Online Games: A Social Network Perspective. ACM Transactions on Internet Technology. (2014).

Jaynes, E.T. 1957. Information Theory and Statistical Mechanics. Phys. Rev. (1957), 620–630.

KasperskyLab. Number of the year: Kaspersky Lab is detecting 315,000 new malicious files every day. http://www.kaspersky.com/about/news/virus/2013/number-of-the-year (2014)

Landauer:, R. 1961. Irreversibility and heat generation in the computing process. IBM Journal of Research and Development. 5, (1961), 183–191.

Lloyd, S. 2007. Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos. Vintage. (2007), 256.

Machta, J. 1999. Entropy, information, and computation. Am. J. Phys. 67, 121 (Dec. 1999).

Maiberg, E. 2014. Hacks! An investigation into aimbot dealers, wallhack users, and the million-dollar business of video game cheating. PCGamer. (2014).

PandaLab. Annual Report Pandalabs 2013 summary (2014). http://press.pandasecurity.com/

Plenio, M.B. and Vitelli, V. 2001. The physics of forgetting: Landauer's erasure principle and information theory. Contemporary Physics. 42, 1 (2001), 25–60.

Plumer, B. 2012. The economics of video games. The Washington Post. (2012).

Reeb, D. and Wolf, M.M. 1962. (Im-)Proving Landauer's Principle. arXiv:1306.4352. (1962).

Scarrott, G.G. 1986. The Nature of Information. Computer Journal. 32, 3 (1986), 262–266.

Shannon, C.E. 1993. Collected Papers. Wiley-IEEE press ISBN-. (1993), 1–968.

Shannon, C.E. 1948. The mathematical theory of Communication. Bell System Technical Journal. 27, 1 (1948), 623–656.

Shannon, C.E. and Weaver, W. 1949. The Mathematical Theory of Communication. Univ of Illinois Press. (1949).

Spiro, T.G. and Stigliani, W.M. 1980. Environmental Issues in Chemical Perspective. Suny Press.

Zuse, K. 1969. Rechnender Raum. Friedrich Vieweg & Sohn.

Zuse, K. 1967. Rechnender Raum. Elektronische Datenverarbeitung. (1967), 336–344.

Wiener, N. 1948. Cybernetics or Control and Communication in the Animal and the Machine. MIT Press.

Williams, M.A. 2008. Representation = Grounded Information. PRICAI. (2008), 473–484.