Liability Transfers in Network Exchanges

Raphael Yahalom

Onaro Research *

yahalom@onaro.com

Workshop on Economics and Information Security - UC Berkeley 2002

Position Paper

May 8th, 2002

1 Information Asymmetry in Network Exchanges

According to the Varian-Anderson principle [8, 1] liability should be assigned in a manner that ensures that "*those who are best positioned to control the risks have appropriate incentives to do so*" (Varian in [8]). As stated by Anderson in [1] "*Where the party who is in a position to protect a system is not the party who would suffer the results of security failures, then problems may be expected*".

However this principle can be extended to also cover cases in which parties are well positioned, and have incentives, to *maliciously perpetrate certain risk events.* Where applicable, it is desirable to try and assign appropriate liability to such players.

Consider an environment in which 2 agents P and Q are using their computing devices to communicate with each other by exchanging messages over some network.

In such cases, for every message m there is no upper-bound on the interval between the events $E_{send(m)}$ at P and $E_{receive(m)}$ at Q.

That is, although most messages sent by P are expected to be received at Q within a short period of time, some will incur significant delays, and a few may occasionally be lost before

* The author is performing part of this research work at MIT - Sloan School of Management

being received. Such message delays and losses may be due to any one of a number of reasons including congestion, transmission failures, component failures, non-malicious errors, malicious attacks by external perpetrators, or deliberate actions by Q.

Thus P cannot be certain that Q has already received m until after it received some consequent message back from Q acknowledging the receipt of m. We consider such acknowledgement message to contain (explicitly or implicitly) an *information commitment* by Q, representing Q's statement about the time at which the information represented by m became available to him.

However Q may lie in his information commitment and claim that he received m later than he really did (or not at all), without P necessarily being able to dispute such a lie.

We refer to the period of time between the point in which Q first received m (and so obtained the information represented by m) and the point in which Q claimed he received m (in his information commitment) as the *pre-committed information interval.* We note that the length of the *pre-committed information interval* itself is unbounded.

2 Exploitation of Information Asymmetry

Depending on a-priori application contractual agreements between the agents, there are a variety of cases in which for a given message m from an agent P, pre-committed information intervals of different lengths correspond to different levels of utility for Q.

In some such cases Q may thus be expect to exploit this type of information asymmetry, as well as the fact that the length of the pre-committed information interval is under his full control, to select a length that would maximize his utility. In particular, a rational strategy by Q following receipt of a message m may be to analyze the information represented by m, possibly to generate some additional events during the pre-commit information interval, and only at some later point to select an optimal information-commitment time and generate a corresponding message.

On the other hand, Q's selection of the pre-committed information interval length may often be sub-optimal with respect to P's utility. However P has no way to control the length of that interval, or to know or prove that Q is behaving opportunistically.

There are examples of exploitation and potential exploitation of such information asymmetry in various applications. Q may be an electronic election service which either discriminates against certain voters based on demographically-based assessments of their likely votes (e.g. [9]), or simply waits until after the election results are published to determine whether to nullify them by making additional information-commitment claims (a very similar highly publicized event may have happened at the *Vivendi* share-holders meeting in Paris last week [3]).

Q may be an electronic trading service that discriminates against certain investment requests and potentially delays their execution in order to exploit market fluctuations or to perform additional opportunistic trades in advance of the requested ones.

Q may be any business entity that in general wishes to defer (in some cases indefinitely) his acceptance of certain liabilities, such as key-revocation requests, transaction-cancellation requests, etc.

Conventional approaches for establishing *happened-before* relationships between events in distributed computing systems (denoted $E_i \rightarrow E_j$), first introduced in Lamport's seminal paper [4], cannot necessarily be relied upon for these type of cases.

As an example, consider P and Q to be traders in 2 different banks conducting derivative transactions by sending order statements that are digitally signed with their respective private keys, (e.g. [5, 7]). Assume that due to some rogue-trader activities discovered at P's environment, P sends an urgent request $\{\ldots$ *revoke key* K_{P-} $\}$. Such a private key request represent P's request to transfer from itself the liability it has on any statement digitally signed by key K_{P-} . Now assume that shortly *after* it receives the revocation request, Q receives a request signed with the key K_{P-} $\{\ldots$ *sell options x* $\}_{K_{P-}}$.

There are circumstances in which Q may have an incentive to go ahead and perform the transaction, and later successfully claim that in his node $E_{sell.options.x} \rightarrow E_{revoke.K_{P-}}$. Note that conventional use of timestamps in the respective messages does not necessarily help because the point in time in which Q accepts the key revocation request is, within certain constraints, under Q's control. Furthermore, a possible contract in which liability would have automatically been assigned to Q within some pre-specified interval from the timestamp generated by P, would have enabled P to lie opportunistically without being detected (for example by generating older timestamps in certain messages).

The economic exploitation of this kind of information asymmetry is not new and is not unique to electronic information exchanges.

However we claim that their overall significance is likely to increase as more business and social exchanges are conducted remotely using networked environments. This informal claim is based on the following general arguments:

- Because of P's limited information about Q's internal environment and about the network transmission environments, Q is in a position to extend the length of the pre-committed information interval with minimal risk to its reputation (e.g. Q can make plausible claims justifying delays, without P being able to verify these);
- Because of the computing power available to him, Q can use the information of m to compute rapidly the optimal information commitment point (e.g. which requests to discriminate against);
- Because of the efficient market dynamics, Q may exploit even short uncommitted-information intervals for his own advantage (e.g. by rapidly taking actions which exploit his early access to the new information);

We note that for similar reasons, with respect to certain interactions, Q may not have a high incentive to invest in infrastructure that maximizes his ability to handle requests rapidly and reliably. The more uncertainty about the reliability and performance (of Q and the network environments) the better Q may exploit the information asymmetry without significantly damaging his reputation. For example, in the recent *Vivendi* events last week [3], *Vivendi*'s management claimed, after share-holders electronic elections results which were unfavorable to it were published, that a computer security breach in *Vivendi*'s electronic elections prevented certain valid voters from recording their correct votes, and so that the elections results should be nullified.

3 A Moral Hazard Class Problem

Q's capability and incentive to extend the pre-committed information interval (possibly indefinitely) and to exploit his possession of the uncommitted information, is an additional risk for P and a decrease in overall market efficiency. In fact we consider this to be a generic special case, fundamental to interactions over distributed networks, of the moral hazard problem (e.g. [6]).

When and whether a message m has been received becomes the hidden-information of the intended recipient Q. Often the incentives of P and Q are not aligned: P's interest that his message m (e.g. a request) will be processed in a timely manner may not be compatible with Q's interest. Furthermore, Q's actions that affect the speed with which m is processed, are not observable by P.

In general, when according to an a-priori contract between P and Q an information-commitment event by Q on m implies a certain liability assigned to Q, then Q may have the capability and the incentive to deliberately delay, possibly indefinitely, its commitment event after the receipt of m. Even without deliberate intent, Q may not have the incentive to ensure rapid processing of certain incoming messages (e.g. may not have an incentive to invest in sufficiently reliable, highly-available, and secure infrastructure to process these messages).

Obviously as with many moral hazard cases, the incentive to exploit opportunistically the information asymmetry decreases in market situations in which there are repeated interactions, and in which significant reputation loss offsets the potential opportunistic gains.

However the characteristics of the network exchange environments enable increasingly sophisticated forms of the types of *selective denial of service* exploitations we are considering in this paper. Using his computing power, Q may determine rapidly specific instances in which discrimination is particularly attractive. Furthermore Q may perform such discriminations with enough sophistication (in term of frequency and scope) to counter various attempts by external parties to reach significant reputation-damaging conclusions based on statistical analysis of observations.

4 Proposed Solution

Our proposed solution for this generic type of *message-transfer moral hazard* is to separate the information commitment event from the control of the agent to whom the liability is assigned. The a-priori contract between P and Q is defined not necessarily in terms of the information-commitment events at Q, but possibly in-terms of observable (and non-disputable) information-commitment events occurring at a dedicated disinterested 3rd party (which we call CTS - *Certified-Transfer Service*).

One of the key principles in our proposed design of a CTS is that, although it may support the transfer of any message m from P to Q, it may not unilaterally obtain any information about the content of m, or possibly even about the identities of P and Q. That can be achieved by using cryptographic techniques (e.g. P may encrypt m with Q's public key) [5, 7], and anonymous channels (between P, Q and CTS) [2]. CTS generates the information-commitment event by digitally-signing, and posting at a specified index location, the cryptographically-protected blob that was provided to it by P (and which Q may then retrieve and recover m).

The fact that the CTS may obtain no information about any particular message implies that it has no conflicting incentives to selectively discriminate against any particular messages by delaying the information-commitment events for them.

Contracts between P and Q may now include obligations by Q to perform certain actions (or accept certain liabilities) within a specified bounded interval of time after particular information-commitments events generated at CTS. Q has strong incentives to adhere to these obligations because he may be provably accountable for his violations. Furthermore note that not only the actual time of each information-commitment event at the CTS is certified, but also the fact that the blob representing the information is actually available for the recipient. Consequently, with proper design of the application exchange protocols, in case of a violation P may prove that Q is responsible because he violated his contractual obligations, given the information he obtained by a particular point in time.

Q on the other hand has strong incentives to initiate periodically look-ups at the CTS, in order to ensure that he obtains information designated to him in a timely manner that enables him to meet his contractual obligations. In actual executions P will likely initially send a message directly to Q and only in cases in which he does not receive information-commitment response rapidly from Q, P will resend the message via the CTS.

Thus by providing an option to rely on a reliable disinterested 3rd party CTS in distributed networks, the risk for agents such as P is reduced, and the overall market efficiency is increased.

5 Conclusions

There are some interesting challenges associated with this CTS solution, computer-science ones as well as economic ones:

- How to design a highly-available, highly-replicated, and highly-secure dedicated CTS with the required characteristics?
- How to design the application protocols between agents, as well as the contracts between them, so that any violation will lead to provable incrimination of the guilty party without violating protocol goals (e.g. maintaining participants' confidentiality)?
- How to analyze the market dynamics when a CTS is introduced?

These challenges and others are currently the focus of our on-going work.

To conclude, electronic network environments are expected on one hand to increase the potential scope of certain types of *message-transfer moral hazard* problems, but on the other hand these environments also enable advanced approaches, such a *Certified-Transfer Service*, for addressing them.

References

[1] R. Anderson "Why Information Security is Hard - An Economic Perspective", *Proceedings of the 17th Computer Security Applications Conference*, New Orleans, Louisiana, Dec 2001.

[2] D. Chaum, "Security without Identification: Transaction Systems to Make Big Brother Obsolete", *Communication of the ACM*, Vol 28, No 10, pp. 1030-1044 Oct 1985.

[3] CNN, "Experts: Vivendi Hack Charge Doesn't Compute" *CNN April 30th, 2002*, *http://www.cnn.com/2002/TECH/internet/04/30/vivendi.vote.hack.reut/index.html*

[4] L. Lamport "Time, Clocks, and the Ordering of Events in Distributed Systems" , *Communication of the ACM*, Vol 21, Num 7, July 1977, pp. 558-565.

[5] R. Rivest, A. Shamir, and L. Adleman "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems", *Communication of the ACM*, Vol 21, No 2, pp. 120-126, Feb 1978.

[6] B. Salanie, "The Economics of Contracts", *MIT Press*, 1997.

[7] B. Schneier, "Applied Cryptography", *2nd Edition, John Wiley & Sons*, 1996.

[8] H. Varian "Managing On-Line Security Risks" , *Economic Science Column, The New York Times*, June 1 2000, *http://www.nytimes.com/ library/financial/columns/060100econ-scene.html*

[9] R. Yahalom "Delegated Enforcement in Secure Protocols" , *2nd Workshop on Issues in the Theory of Security*, Portland, Oregon, Jan 2002, *http://www.dsi.unive.it/IFIPWG1_7/WITS2002/prog/annotated_program.html*