— Impersonation in Electronic Commerce —
“Legal Rules for Allocating Risks of Impersonation and Repudiation by Sellers in Commercial Transactions on the Internet”

This thesis is submitted in partial fulfillment of the requirements for the Honours degree of Bachelor of Laws in Monash University Research Subject 512
(Course Code: LAW 5207), Faculty of Law
Monash University
September 1998

By G. Brendan Macmillan
Abstract

"Should you be bound by your word even if you did not say it?"

Yes, according to the recipient-based liability of credit cards, provided that the word appeared to be yours from the recipient’s point of view. This is contradicted by a resounding No from the traditional view that you cannot be liable for impersonations.

The ability to rely on appearances is very attractive for internet commerce. Such commerce is electronic and takes place at a distance, where there is no physical presence and it is often difficult to authenticate the sender’s identity. Enabling recipients to rely on an appearance, without bearing the risk of impersonation, could be the key to the internet marketplace.

Internet commerce is potentially highly lucrative. The internet is faster, cheaper and further reaching than any other medium; it could overthrow the tyranny of distance once and for all, and open the world up for commerce even more than water-carriage ever did. But a lack of certainty and safety for recipients will stifle its growth. This thesis argues that recipient-based liability is possible, but only if each reliance on your word is registered - just as are your credit card transactions.

The internet: electronic, distant, broad

The internet is an electronic medium, far-reaching to many and distant points on the globe. This presents problems: the problem of electronic authentication at a distance is that personal contact and other means of authentication are generally not available. The breadth of the internet is a problem because many people may be deceived by an incorrect authentication: since the potential number of recipients may be in the millions or billions, this factor is not inconsequential.

Electronic Because the internet is electronic in nature, traditional tokens and paper-based methods of authentication are not available. A piece of paper cannot be transmitted digitally - only a representation of it can be. The same is true
of biological identifiers (such as fingerprints and handwritten signatures). In this
digital form, they lose their status as tokens and can only facilitate authentication
if they are kept secret.

However, the inappropriateness of tokens in an electronic medium largely has
been solved with encryption, which involves a secret encryption key. Addresses are
another means of authentication as a message sent to your address may be presumed
to have been received by you.

Chapter 2 discusses these authentication mechanisms of tokens, secrets and ad-
dresses in greater depth.

Distance One of the great advantages of the internet is that distant locations can
be accessed (almost) as easily as those that are local. A serious problem with this
is that it becomes very difficult to authenticate the origin1 of messages. There is no
physical presence, and little in the way of surrounding circumstances: impersonation
is easy on the internet. Recipients would be reluctant to rely on such messages, thus
stifling commerce.

To combat this, there have been several proposals for recipients of an “electronic
signature” to rely on it completely, even if it turned out to have been sent by an im-
poser. One such proposal is article 13 of the UNCITRAL Model Law, the subject of
Chapter 4. This would be convenient and safe for recipients, who need only consider
the appearance of the signature, and not seek alternative means of identification.
This shifts the risk of impersonation from the recipient to the purported sender.

Broad The internet is possibly the broadest means of communication that the
world has ever seen. It is global and instant. This makes it an “open environment”
in the sense that one may communicate with any number of people and contract
with each of them.

Unfortunately, this large number of potential recipients renders the above allo-
cation of risk infeasible.

Allocating Risks

“Should you be bound by your signature if you did not sign it?”

5Sender-based Liability Consider the traditional liability rule for signatures2 that
you are not liable for forgeries. If you signed then you are bound; but if you did not
sign, then you are not bound - it is based on your point of view. Even if the forgery

1Chapter 1 argues that identification is essential for trust, which is essential for commerce,
though often the identification need only be that the party is the same as that to a previous course
of dealing, or that the person selling the gold is the same person who owns it.

2We focus on the identification function of signatures; not as indicators of a special kind of as-
sent, nor their ritual, ceremonial or certainty roles. Note that in commerce, actual “identification”
is not always required: you want to be sure that the person selling you the gold is the same person
who owns that gold; but you do not necessarily care if that person is Joe Bloggs or John Smith.
is excellent and indistinguishable from your own, you are not bound, provided you can prove this. Although an excellent forgery might fool a handwriting expert and thus a court in its forensic examination, you can lead evidence beyond the signature itself to show that it must have been forged: you were in Peru on that day with two broken arms. Such evidence is of course not limited to what the recipient could have known at the time of receipt.

For example, a rogue impersonates me and purports to sell my gold to a third party. The third party pays for the gold and the rogue promises delivery under a contract of sale - in my name. However, by the forgery-rule, this contract does not bind me. The rogue is happy; the third party is unhappy; and I am sanguine.

The third party learns their lesson: they rely on a signature at their peril, and so are motivated to seek other indicia of authenticity. They take the decision of whether to rely on the signature or not, and also bear the risk of their being mistaken.

**Recipient-based liability** Consider an alternative rule by which you are liable for forgeries. Such a rule would be attractive for the electronic transactions at a distance of the internet because it would provide certainty for recipients.

The scenario now plays out quite differently: the rogue again pretends to be me and enters a contract to sell my gold to a third party. This time, the third party knows that if the signature looks authentic at the time of signing, a court will uphold the contract and it will be enforceable against me - whether I was impersonated or not. The rogue again makes off with the money; but when the third party seeks to enforce the contract against me and I complain that I was impersonated, the court replies:

“We don’t care if you didn’t sign it. We don’t care if you were in Peru with your arms broken. We don’t care if it was a forgery. It appears to be your signature; you are liable”.

The rogue is happy. The third party is happy. I am unhappy.

What can I do about this situation? I could make my signature more difficult to forge, using very complex and intricate sworls, be trained by a handwriting expert, and use a quill pen so that each change in pressure of my hand is registered on the paper. I could make it very difficult to forge. And let us assume that I could say to the world that my signature is not to be relied upon over the value of $50. Now, a rogue comes along and thinks: “It’ll take me weeks just to learn this guy’s signature. A lot of work! And for a measly $50? It’s just not worth it!” By making the signature more difficult to forge, and by reducing the rewards, it seems we can make forgery a non-issue.

Assuming it is my responsibility to make my signature difficult to forge, I am arguably at fault if a rogue does forge it, because I must not have gone to enough trouble. Recipient-based liability may begin to seem reasonable. However, these rules apply in an *open environment*, where the signature is good to the whole world. A rogue impersonating me could enter a transaction with one person for $50, and
with someone else for another $50, and with every person in the country or in the world. The total liability might be $50 million or $50 billion. It is difficult to determine just what the potential liability would be, but it could be large: the rogue reconsiders, and decides that $50 billion is worth a few week’s work after all.

What could I do in the light of this crushing liability? I could make my signature even harder to forge, so that it is so difficult that it is not worth the trouble for even $50 billion. It is hard to imagine a signature that is that difficult to forge. It comes down to a question of balancing the difficulty of impersonation against the potential reward to a rogue; and that the breadth of the internet multiplies this reward to many time larger than it might at first appear to be. The adequacy of security must be weighed against not just the probability of failure; but also the consequences.

And that is my first point: that recipient-based liability rules cannot work in an open environment – without something more.

Existing recipient-based liability
An extraordinary but necessary incident of this recipient-based approach is that the recipient could sometimes be able to rely on a message that only *appeared* to originate from the sender, despite a court finding in reality that it did not so originate. That is, the sender could be held to their signature *even though they did not sign it*. But this extraordinary result is not without precedent.

An EFTPOS card is an example of such a recipient-based scheme: withdrawal or payments by a rogue in possession of the card and PIN are deemed to be from the owner\(^3\), in that the owner’s bank account is debited, and the owner has no right as against the bank to reverse the transaction\(^4\). Evidence that the transaction was not entered into by the owner of the card is to no avail.

The significance of this is not that the purported sender is liable; but that the recipient is not liable. In the credit card scenario, the merchant is paid if the transaction appeared to be authorised. Thus, the merchant is happy to rely on payment by credit card. Where the cardholder is not liable for the whole or part amount, the credit card company pays the difference, who effectively insures against this risk. However, this risk is only insurable if the total extent of the risk - in terms of probability and size - is determinable.

This thesis argues that the only way for a recipient-based scheme to be workable is for liability to be limited in some way, just as it is with credit cards and EFTPOS cards. Such *loss limitation devices* are discussed in section 3.2. Limitation of actual losses allows a party’s legal liability to be limited - but without it, such reallocation is merely an exercise in rearranging deckchairs.

It is argued that one reason for the forgery rule for signatures is because they

\(^{3}\)Other rules contained in the EFT Code of Conduct soften this basic allocation of risk, having regard to whether the card was reported as stolen and so on; but the basic rule may be construed as stated.

\(^{4}\)Of course, the owner has a cause of action as against the rogue; but in practice this is not very helpful, as the rogue usually is long gone.
also operate in an open environment. Without such a rule, an imposter could bind you to any number of contracts.

Registration of reliance

The problem of recipient-based liability is the open environment. It is possible to provide recipients with the convenience and certainty of a pin-point of identification, simply by closing the environment. This may be done by reducing the number of potential recipients - as in EDI\textsuperscript{5} - or it may be done using registration.

The idea of registration is to prevent multiple (or overlapping) reliances. It is an old idea that is reused again and again in various legal regimes, such as in credit card transactions and the Torrens land system. The paramount purpose of these systems is that registration reflects the true state of affairs, so that the register may be relied upon as accurate by other parties, thus providing a single point of information whose appearance reflects reality. This provides convenience and certainty for recipients. Without a registration system, multiple-selling of land \textsuperscript{6} or multiple-spending of money is possible; for identification on the internet, it is the generalised danger of multiple-reliances.

Conclusion

Chapter 5 brings together the ideas discussed throughout the thesis to form a principled basis for allocating risks of impersonation in internet commerce.

\textsuperscript{5}Electronic Data Interchange - where two parties agree to exchange data electronically. Typically, this data comprise orders that are legally enforceable. For example, a large department store may automate its inventory control, so that orders are automatically placed with supplies when stocks levels fall beneath a threshold.

\textsuperscript{6}or of overlapping interests, such as multiple mortgages etc
Acknowledgments

I would like to thank Mark Sneddon for asking exactly the right questions (it has taken me over two years, but I think I may have finally answered them), Nicholas Pengelley for his sage advice and rather astonishing ability to conjure up research articles, Shane Bruning, Sky Christensen, Michael Wilkinson for heroic stamina through extended discussions, Michael Attipa for an insight that caused me a lot of thought, Chris Betts and Tony Jansen for their embarassingly perceptive commentary, advice and proofreading, Ian Macmillan for his challenging questions and warm encouragement and the many, many with whom I have discussed the ideas in this thesis, including (in no particular order) Jamie Scuglia, Lloyd Allison, Bruce Dyer, Chris Wallace, Rob Rendell, Jane Brion, Prof. McCredie, Jeff Goldsworthy, Jeannie Paterson, Elizabeth Lanyon, John Glover and Jaqui Lipton.

I gratefully acknowledge Melissa de Zwart’s supervision of this thesis and I also wish to thank the Faculty of Law and the Department of Computer Science and Software Engineering of Monash University, which have facilitated this thesis in many ways, and whose academics, staff and students have made available their intellectual power to this quest; indeed, several of the above persons are members of these bodies.
Contents

1 Electronic Commerce ........................................ 1
   1.1 The Internet ........................................... 1
   1.2 The Necessity of Trust .................................. 2
   1.3 The Origin of Trust ..................................... 3
   1.4 Identification .......................................... 4
   1.5 Identification Risks ................................... 6

2 Authentication Technology ................................. 7
   2.1 Tokens ................................................... 8
      2.1.1 Imprints ............................................ 8
      2.1.2 Biometrics ......................................... 9
      2.1.3 Tokens cf Information ............................. 10
   2.2 Secrets .................................................. 11
      2.2.1 Choice .............................................. 12
      2.2.2 Transmission ....................................... 12
      2.2.3 Encryption ......................................... 13
      2.2.4 Control cf Intrinsic risk ......................... 15
   2.3 Addresses ............................................... 17
      2.3.1 Callback ........................................... 17
      2.3.2 Internet Addressing ................................ 18
      2.3.3 Forging Return Addresses ......................... 19
   2.4 Conclusion .............................................. 20

3 Administration Risks ........................................ 23
   3.1 Application .............................................. 23
   3.2 Limitation ............................................... 24
      3.2.1 Breadth ............................................ 25
      3.2.2 Independence ...................................... 26
      3.2.3 Registering Reliance ............................... 27
   3.3 Revocation ............................................... 28
   3.4 Conclusion .............................................. 28
## Contents

4 Article 13

4.1 The Operation of Article 13 ............................................ 31
4.1.1 Appearances ................................................................. 32
4.1.2 Control ................................................................. 33
4.1.3 The Notice exception ................................................. 35
4.1.4 The Knowledge exception ........................................... 35
4.1.5 Consequences of Attribution ......................................... 37
4.2 The ECEG position on Article 13 ................................. 37
4.3 Conclusion ................................................................. 39

5 Conclusion ................................................................. 41
Chapter 1

Electronic Commerce

1.1 The Internet

“As by means of water-carriage a more extensive market is opened to every sort of industry than what land-carriage alone can afford it, so it is upon the sea-coast, and along the banks of navigable rivers, that industry of every kind naturally begins to subdivide and improve itself.” — Adam Smith

The information-carriage of the internet may open a yet more extensive market than water-carriage ever did, for the internet is cheaper, faster and connects more broadly.

But just how does the internet differ from existing forms of electronic communication, such as the telephone? “Electronic” commerce is not new; after all, the telegraph is now obsolete technology. The phone, telex and fax also have been with us for some time. All are cheap, fast, convenient and connect broadly over the world. As in the industrial revolution, it is automation that makes the difference.

Ironically, the flow of commerce on the internet is damned by the very things that bless it: while automation opens a wider market; fraud may also be automated and over an equally broad extent of market.

“The net offers Australian business - dogged by the tyranny of distance for two centuries - the chance to turn small domestic markets into large global ones, using the wonders of digital technology to approach customers previously beyond reach [to sell various physical goods].” — Australian Financial Review, Editorial

Unfortunately, the tyranny of distance still dogs the delivery of goods: delivery

---

1An Inquiry into the Nature and Causes of the Wealth of Nations (1776) Book 1, Chapter 3.
2EDI (Electronic Data Interchange) has even automated contract formation since the 70’s.
times for CDs on the internet vary from days to weeks\textsuperscript{4}. However, delivery of information goods and services is possible over the internet itself. CDs are a case in point, for depending on the speed of connection, the information contained in the CD can be delivered in a matter of minutes, not days or weeks, and without the expense, risk and uncertainty of international shipping\textsuperscript{5}. The on-line execution of contracts - delivery and payment\textsuperscript{6} - truly would overthrow the tyranny of distance.

By lowering entry and transaction costs, the extent of market of both buyers and sellers may be increased - and hence the wealth of the nation.

1.2 The Necessity of Trust

Commerce on the internet, as elsewhere, requires trust. Unfortunately, trust on the internet is a real problem that must be solved before we can reap the benefits of information-carriage. This section shows why trust is essential to commerce.

Commerce is the exchange of one thing for another. This corresponds with the commonsense notion of a bargain: for example, you might purchase a peppercorn for a dollar in a sale of goods. A promise of future performance may also form part of an exchange. Trust is a confident expectation or reliance\textsuperscript{7}. To trust someone is to expect them to perform their side of the bargain; and to perform it properly. You might trust someone to keep their promise to deliver the peppercorn; and that the peppercorn is up to standard.

In an atmosphere of mutual trust and respect, the issue of trust is not apparent. You chat with merchant, recounting your adventures on the high seas amongst the aromatic sacks of spices, and the exchange of peppercorn and coin might as well be simultaneous. There is no sense of vulnerability as you relinquish your coin. But when you trade at a distant port, where that respect is stripped away and replaced with mutual suspicion, the role of trust becomes conspicuous by its absence. You are fearful of letting the merchant have your coin until he gives you the peppercorn - he might take the coin, and then claim that you never gave it to him! But the merchant does not trust you either; he thinks that if he gives you the peppercorn, you will run off without paying. Deadlock.

It is tempting to suggest that a simultaneous exchange could resolve this childish dilemma, so that neither party need go first. But is this really possible? Or does

\textsuperscript{4}This irony is well-illustrated by the story of an internet transaction to a room full of impressionable journalists: after completing the “instant” transaction, the demonstrator quickly turned off the application. But not quickly enough to hide the message displayed: “Please allow up to 6 weeks for delivery”.

\textsuperscript{5}Another advantage of information goods is that a claim of non-receipt can be rectified by simply resending the goods, with insignificant cost to the seller.

\textsuperscript{6}The problem of payment on the internet is beyond the scope of this dissertation; but has received much attention, with proposals for electronic money and safe credit card payment (using the proposed Secure Electronic Transactions standard, or SET). Of course, Electronic Fund Transfers (EFT) have been used by banks for many years now.

\textsuperscript{7}combining the first and second definition given in the Shorter Oxford Dictionary
one party necessarily lose control of their consideration before the other? A series
of thought experiments will confirm that trust is always required at some level8. In
addition, the physical presence of simultaneous exchange is not available on the
internet.

If trust is essential for commerce, then the ceaseless exchange of coin and promise
and thing that permeates our lives and creates our wealth also implies tremendous
trust. How is this trust established? Why do we trust some people and not others?
The keys are reputation, rapport and recourse.

1.3 The Origin of Trust

Trust usually requires identification: Little Red Ridinghood trusted her grand-
mother, but this trust was to no avail when her grandmother was impersonated
by a wolf. Sometimes it is not enough to trust someone; but you must also confirm
that it is they who you are dealing with. Trust and identification are distinct issues,
but trust rests on identification.

When poised at the threshold of a transaction, we can look in three directions
for a reason to trust the other: backward to the past, to the present, and forward
to the future.

Past Trust based on the past is quite reliable9 and may be built up over a course
of dealing, the trust reposed increasing with each successive transaction. Unfor-
tunately, past acts can only suggest the nature of the party concerned, without
guaranteeing that they will carry on in the same manner. Their inner nature may
change; or we may have been taken in by a clever act. This might be known as the
"extended courtship" theory of trust.

You prefer to deal with a merchant you know has been trustworthy in the past,
rather than an unknown stranger or someone with a positive reputation for untrust-
worthiness. This type of trust hinges on accurate identification, for your trust is
mislaced if you are fooled by an imposter, as was Little Red Ridinghood.

Other forms of reputation can also establish trust, such as by word of mouth,
recommendation by independent body (eg RACV, Choice Magazine) and the par-
tisan recommendations of advertising, packaging and salesmanship, but again some

8Escrow makes the trust explicit. If you and the merchant both trust me, then I can facilitate
trade between you by acting as an escrow agent. For example, you give your coin to me and the
merchant gives me the peppercorn, and you both trust me to complete the transaction. If the
merchant reneges, and refuses to give me the peppercorn, I simply return your coin with no harm
done. Escrow agents are employed in international shipping (usually banks), with payment and
delivery evidence by irrevocable letters of credit and bills of lading, respectively.

9Trust is "traditionally established over time" Electronic Commerce: Building the Legal Frame-
work - Report of the Electronic Commerce Expert Group to the Attorney General, 31 March 1998,
p. 111, 3.1.2; hereafter referred to as the ECEG report.
form of identification is required to assure us that the party we are dealing with is the same one advertised.

**Present** However, when the recommendation occurs in the present, no identification is required; for example, when the other says “trust me” - and we do. This might be called the “love at first sight” theory of trust. Although instant rapport is possible, it generally takes some time. That is, it is based on the past.

The distinction between the present and past forms of trust is between walking into the merchant’s storehouse for the first time and judging it at face value, and choosing a merchant you have dealt with before who has a good reputation.

**Future** The availability of recourse in the future is the third basis for trust. Recourse can take many forms; we here focus on legal recourse and recourse against the other’s reputation. The hard work of establishing a good reputation leads to a desire to protect this investment, and to provide a means of identification which facilitates consumers relying on this reputation from the past. Thus, the threat of damage to reputation in the future provides a basis for trust today. For recourse to be effective, the other must have something to lose - such as funds or reputation; and some way to lose it - such as through a court or word of mouth; as well as being identifiable. This might be described as the “retribution” theory of trust.

Only trust based entirely on the present needs no identification - but “love at first sight” is the least reliable basis of trust. Trust based on the past and trust based on the future both require some form of identification.

### 1.4 Identification

Identification itself may be weak or strong. Weak identification is where you deal with the same person each time, but you do not know who they actually are. For example, a fellow commuter lets you into traffic. After you have switched lanes, you wave a thank-you. You are certain that it is the same person who lets you in and who you wave at - but you do not know who they are. Strong identification is when you also know who that person is: you stop beside them at traffic lights, look across and recognise an old friend: “It’s you!”

Some of the dangers of weak identification are that multiple personas may be used by the one person, and that legal recourse to the actual person is not available.

**Reputation**

Trademarks function primarily as weak identifiers: you assume the branded product originates from the same source each time you purchase it; but you may not know who, what or where that source is. You are content to repose trust in the reputation
of the trademark. Of course, should you want to discover the actual source, you usually can.

Weak identification has the problem of not necessarily giving a complete picture of the person in question - the Dr. Jeckle and Mr. Hyde problem. All we know is that it is the same party as before; we do not know what else they have been up to. The one company might produce a range of “environmentally friendly” products under one label, and “environmentally unfriendly” products under another. Knowing about one persona does not necessarily tell you that much about the person behind it. The throw away identity is another possibility, where untrustworthy acts are carried out under an alternate persona, which is then discarded.

Despite this, trademarks do seem to be reliable in general. This is because a good reputation is worth protecting; and a substantial investment in reputation is too valuable to throw away. Identification in the future is the basis of this trust; and an identification that need only be weak.

That a person may establish more than one such identity or persona is little different from one company holding many different trademarks. Though there may be an element of Dr Jeckle and Mr Hyde in these brands, the desire to protect the investment in reputation seems adequate to establish trustworthiness.

Legal recourse

Legal recourse would seem to demand strong identification of the actual party.

However, judgement in a civil case is usually in the form of damages, suggesting that a fund\textsuperscript{10} associated with the identity might be sufficient to satisfy judgement. This is similar to the present situation with companies, where only the company itself is liable for judgements against it, as an artificial person, rather than the management or shareholders. A problem occurs where the liability exceeds the fund, or where the gravity of a crime committed by the company requires that the persons behind it be punished. A fund limited to the purchase price of a transaction would assure customers that at least they could get their money back; but may not be sufficient to meet the requirements of justice and regulation. Note that recourse against this fund would require access to a legal entity that is strongly identified.

On the other hand, being unable to meet liabilities is a familiar scenario\textsuperscript{11} - the “limited liability” company explicitly shields shareholders from such liability. The sometime inadequacy of the proposed fund would simply be more of the same. Weak identification is not as good as strong identification, but it may be sufficient to provide the trust necessary for commerce.

In conclusion, it seems that reputation coupled with legal recourse (against a fund of money) could engender the trust necessary for commerce; but justice would be frustrated in some cases because regulation and the protection of the public good

\textsuperscript{10}This fund could be provided by an insurer.

\textsuperscript{11}This scenario is so common that specific bankruptcy and insolvency laws have been developed to handle it.
require strong identification. At any rate, it is suggested that legal recourse acts mainly as a safety net; reputation is the mainspring of trust. Weak identification is sufficient for commerce.

1.5 Identification Risks

Identification is essential for reputation and recourse - for without knowing to whom it relates, reputation is meaningless; and without a target, recourse (legal or otherwise) is absurd. Identification is essential for internet commerce, as identification brings into play both reputation and recourse.

Impersonation is an obvious danger of unreliable identification; but where impersonation is plausible, there is also the risk that a sender might claim to have been impersonated, and attempt to repudiate their message.

An inked signature stamp belonging to X illustrates three forms of unreliability: a rogue could steal the stamp from X; the stamp itself might be counterfeited by inference from examples of its use; and thirdly, the signature stamp might not really belong to X in the first place. These risks are named the control, intrinsic and application risks.

Control risk The control risk is that control may be lost of the identifying mechanism. Note that the person holding the device is most able to prevent this.

Intrinsic risk The intrinsic risk is a risk that neither the holder nor the recipient can prevent - it is an inherent weakness in the identification mechanism itself. However, note that the party choosing the particular device does have some influence, for some devices are more easily counterfeited than others. Even within the one family of devices, there is variation; for example, an intricate and fine signature stamp is more difficult to counterfeit than one that is simple and rough.

Application risk The application risk is named after the risk in credit cards that the applicant used another’s name when filling out the application form. This is not a risk of the identification mechanism itself, but of its administration.

The control and intrinsic risks are discussed in the next chapter on Authentication Technology; while the application risk is covered briefly in the chapter on Administration Risks.
Chapter 2

Authentication Technology

The Electronic Commerce Expert Group warns of the unpredictability of technology and markets\(^1\), and so urges a principled and technology-neutral treatment of electronic commerce.

We therefore analyse the risks of authentication mechanisms into control and intrinsic risks. We apply this analysis across the board of authentication technologies to show that this approach has universal application.

Although this analysis of risk suggests a prima facie allocation of liability, we focus on the risks themselves in this chapter.

2.1 Possession of a Token

Tokens are physical objects that evidence something through possession: land and share certificates evidence ownership of particular interests in property; possession of a bank book imply an authorisation to deal with the funds in that account. Some tokens go further than proving a specific interest or authorisation: they evidence identity\(^2\). Such tokens include passports, birth certificates and letterheads.

That a rogue might steal the token is the obvious danger, and the level of this control risk is within the continuous control of the token's owner.

However, a rogue might also 'obtain' the token by forging it, which is the intrinsic risk of tokens. The anti-counterfeit devices of currency of watermarks, embedded wires, holograms and fine detailing serve to decrease this intrinsic risk.

Different tokens can have different intrinsic risks, depending on the anti-counterfeit measures employed: the party choosing the token also chooses its intrinsic security. However, once this choice is made, the intrinsic risk is out of the control of the chooser. In contrast, continuous care must be exercised with respect to the control risk.

---


\(^2\)Whether this evidence is conclusive or merely rebuttable is a legal question rather than a technological one, and will be dealt with later.
If the intrinsic security is sufficiently high, it becomes easier for a rogue to steal the token than to copy it, and the control risk becomes the point of vulnerability. For example, currency is more often stolen than forged.

2.1.1 Evidence by Imprint

Although possession of the token itself is the issue, this possession can be in turn evidenced by its imprint; such as the signature stamped on a document, and the wax seal applied to the envelope.

Because an imprint is an extra step away from the token itself, it may be easier to forge than the token itself, as there may be more to the token than the imprint reveals. For example, the ornate handle of a seal does not affect its imprint, and so need not be duplicated if forging the imprint. This secondary risk is still an intrinsic risk, and may be conceptually grouped with the primary risk of counterfeiting.

An imprint does more than evidence possession, for it can also link possession with a document, which is very convenient for evidence.

2.1.2 Biometrics

Handwriting is an imprint of a token of a different kind: it represent the quality of the writing hand - a biometric or “measure of organic life”. The old-fashioned quill pen leaves a more revealing imprint than today’s ball-point pens, because it is more sensitive to the pressure applied, and so measures the quality of the hand more accurately. The “PenOp” signature device referred to by Wright in the Eggs in Baskets article (n3), also measures the speed of motion and changes in pressure of the signing pen and so captures more of the token in the imprint.

Other biometrics include forensic identification by fingerprints, DNA, dental records and so on, and our natural measures of a person’s face, gait, and tone of voice that form an important part of how we ultimately identify one another.

Biometrics are a unique class of identification technology in that they tie directly to the physical person: the control risk is low. However, the possibility of loss of control exists, since an involuntary signing may be seen as a loss of control over one’s self; and the theft of someone’s fingertips is a remote, though gruesome, possibility.

2.1.3 Tokens cf Secrets

Handwriting is only part of the security of a signature: a secret signature is even more secure. However, this choice is information, and not a physical token.

\(^3\)However, note that a thief can steal a particular dollar bill only once; but a counterfeiter is not limited in this way.

\(^4\)Unfortunately, this approach is fundamentally flawed, because the imprint is recorded as information rather than physically.
2.2 SECRETS

Tokens are physical; they cannot be conveyed electronically. To send a signed document in a sealed envelope over the internet is a nonsense confusion of categories. Once a biometric is recorded electronically, it becomes information\(^5\) and information cannot identify as a token; however, information can identify as a secret.

2.2 Secrets

A secret, such as a password, can identify you if you alone (and the gatekeeper) know it - it has to be secret. A gatekeeper only allows entry to those who know the password, thus identifying them. The obvious danger is that someone else may learn the secret from you, and therefore be able to impersonate you - this is the control risk. Less obvious is the risk that the password is guessed, which is the intrinsic risk of secrets. An important aspect of the intrinsic risk is the number of guesses that the gatekeeper will allow.

Tokens often incorporate secrets. For example, a rogue who never sees a sample of your signature would find it difficult to forge. Thus, a confidential signature combines a biometric with a password.

A Personal Identification Number (PIN) is a form of password. The control risk relates to someone observing your PIN when you type it in, or breaking into your home and finding where you have recorded it. That is, learning a secret. If you are sufficiently careful, you can reduce this risk to practically zero; but if you are careless, the risk can be dangerously high.

The same is true of the password you type into a terminal to login to a computer: someone may observe it, or discover where you have recorded it.

However, once the PIN is entered into the Automatic Teller Machine (ATM), or a password is entered into a terminal, the secret is in the hands of the other party. If they are careless, an eavesdropper\(^6\) may learn the secret from them; that is, they may lose control of the secret. This is a control risk, because the secret is not guessed or inferred, but learnt; and because the party with control could have prevented it by exercising greater care. In fact, both parties have control, and both are exposed to the control risk.

In contrast, once a password has been chosen, the influence of either party to prevent a rogue guessing it is curtailed. There is no continuous influence, as for the control risk. That is, the distinction is not a matter of degree (as the words ‘influence’ and ‘control’ might suggest), but of when the party can affect the risk.

\(^5\)Simson Garfinkel with Gene Spafford, *Security and Commerce* (O’Reilly, USA, 1997)

\(^6\)Eavesdropping might operate by someone wire-tapping the terminal or ATM.
2.2.1 Choice

The intrinsic risk of a rogue guessing or inferring\(^7\) a password varies with the specific password chosen. An alpha-numeric password can cover a great many more possibilities than a PIN - though in practice, many people choose easy-to-guess passwords (because they are also more memorable). A password system embraces a family of identification mechanisms - each particular password is a specific identification mechanism. Each has its own intrinsic risk - a password may be short or long, difficult or easy to guess or infer. The password:

“hello”

has a high intrinsic risk; whereas the password:

“5eq8c1snfih297eiNheulr2MG6e5t9y3r6r52wmGFDFkuenbueb”

is much harder to guess, and so has a lower intrinsic risk.

Thus, the party choosing the password also chooses its degree of intrinsic risk. This party influences the likelihood of a rogue guessing the password; but this influence is exhausted once the choice is made. In contrast, the party with control has a continuing choice of how to exercise it.

2.2.2 Transmission

A particular type of control risk for secrets is that they are overheard when being told (or transmitted) to the gatekeeper; in contrast to being learnt from an aide de memoire. While the spoken password traverses the distance to the gatekeeper, it is arguably not within the control of either party. But the speaker has the choice of how and when to say it. If the recipient requires a certain method of transmission (eg via email, or entry on the keypad of an ATM) then they are dictating the other’s exercise of control, and so the control is shared between the parties.

The control risk can be great, if the PIN or password is visible to the world. An example is credit card numbers being transmitted over the internet; however, credit card numbers were not specifically conceived of as ‘secrets’.

Credit Card Numbers

Credit card numbers can be used without the physical card\(^8\); for example, for long distance telephone calls. Unlike the physical card itself - a token - the credit card number can be transmitted on the internet; but as with passwords, it can only identify if it is kept secret.

---

\(^7\)An empirical inference is more often a kind of guess than a certainty - it is an “educated guess”.

\(^8\)The cardholder’s name, expiry date of card and perhaps address may also be required in addition to the number.
2.2. SECRETS

Much has been made of the transmission risk of credit card numbers. It is possible to write a program to monitor traffic through a node on the internet, and capture any credit card numbers passing through. This suggests that posting your credit card number through the internet is as secure as writing it on the back of a postcard. The counterargument is that credit card numbers can already be obtained quite easily, such as by going through the rubbish bins behind a restaurant for discarded credit card receipts.

This analysis of credit card numbers as secrets suggests that they might also be obtained by guessing - theoretically, at least. The intrinsic risk for credit card numbers is that they could be guessed instead of learning them - although some regularities allow inferences which can increase the chances of guessing correctly.

These flaws make credit card numbers an unsuitable mechanism for identification on the internet, in themselves.

2.2.3 Encryption

Encryption is where a key encrypts a message, and transforms it into a garbled (or "cryptic") form. In symmetric encryption, the same key decrypts and ungarbles it. Although encryption is named for keeping secrets, it can also act to authenticate origin: as with a password, only the possessor of the encryption key could have sent the message. This is the idea of the digital signature.

The intrinsic risk is how difficult it is to infer or guess the key from the encrypted data. As with passwords, the longer the key, the harder it is to guess. When an encryption key is stored on digital media, there is no need to make it memorable (such as "hello"), so the choice can be perfectly random, and so more difficult to guess. However, computational power and mathematical insight can be surprisingly effective in making educated guesses, or inferring the key from samples of encrypted messages.

Encryption largely solves the transmission problem, because the key - the secret - is not exposed directly to the world, and so cannot be learnt directly in this way. It is exposed only indirectly, through its garbling of a document. However, transmission still constitutes a partial exposure, for each example of a key's use provides more

---

9A credit card number consists of 16 digits, some of these are limited in their values because of they indicate the branding of the card. The final digit is a check-sum, that is determined by the other digits (This is used to catch transcription errors. The algorithm is fairly simple and publically available.). This leaves 16-4-1=11 digits. Although the chance of guessing a specific card number is a million million to one, the millions of issued cards means that the chance of guessing a valid number is closer to 100,000 to 1. Given many guesses, this is not very safe. Note that the expiry date of the card is usually in effect combined with the number - but the values that a valid card can take are fairly narrowly restricted (perhaps to one month in two years, or 24 to 1). Simson Garfinkel with Gene Spafford, Security and Commerce (O'Reilly, USA, 1997)

10The technological details of implementing encryption, such as message digests, are not relevant to this discussion, so are omitted.

11The length of a key is usually expressed in number of bits - binary digits.
evidence with which to infer it.

This indirect exposure is similar to the *imprint* of a signature stamp, in that the interaction of the key with the message associates the two, providing a similar evidential link from possession of the key to the message, thus authenticating it. However, a serious deficiency for legal recourse is that the recipient of the message could have forged it perfectly, because the recipient also has the secret key\(^{12}\). This shared possession also doubles the control risk, as with passwords, because control of the key could be lost at either end.

The above describes a specific type of encryption, called *symmetric encryption*, so-called to emphasize use of the *same* key for encryption and decryption; and to contrast it with *asymmetric encryption*.

**Asymmetric Encryption** In asymmetric encryption, one key encrypts the message; but a different key is required to decrypt it. This is also known as *public key* encryption, for its use in hiding messages: the encrypting key is made public, while the decryption key is kept private. Only the holder of the private key can read the message.

For authentication, the keys are reversed: the encrypting key is kept private, while the decryption key is made public. When the holder sends an encrypted message, anyone with the public key can determine that a message was sent by that holder - provided that it was kept secret, and a rogue did not guess it. The risks of sharing a key - recipient forgery and doubled control risk - are thus eliminated.

The intrinsic risk of encryption keys being inferred from encrypted messages\(^{13}\) is affected by computational power and the unpredictable advances in mathematics. Moore's law of exponential growth seems to hold - that the speed of computation doubles every 1.5 years. The unbreakable code of today may be trivial to break tomorrow - and inevitably faster to do so, even without discovery of a fatal flaw. The surprising twists of mathematical insight could suddenly leave the impenetrable depths of an encryption algorithm high and dry, exposed for all the world to see.

An alternative avoids completely these intrinsic mathematical and computational risks; but at the cost of sharing keys. It is called the *one-time pad*.

**One-time pad** A one-time pad is a list of keys, each of which is used only ‘one-time’. In effect, it contains a different symmetric key for each character to be encoded. Thus, inferring the key for one part of a message does not help with the next part of that message, or the next message. No matter how many examples of its use a rogue obtains, they will not help with the next message, because there is no relationship between successive keys. Although it cannot be inferred, it theoretically

---


\(^{13}\)Note: knowing the public key of the private/public pair provides extra information that may aid the inference task.
2.2. SECRETS

could be guessed; but since each key is never used again, such a guess could never be verified. The intrinsic risk is zero in practice.

Because the pad is shared between sender and recipient, this is a form of symmetric encryption, with its attendant vulnerability to recipient forgery and loss of control by either party.

A variant uses asymmetric key where both encryption and decryption keys are kept secret: a public decryption key would provide a foothold for inference of the other key. However, the recipient could use this decryption key in this way, and thus raise the spectre of recipient forgery. But clearly, this is an improvement over the recipient having complete control over the secret key, as in the pure symmetric version. Unfortunately, having different asymmetric keys of sufficient length for each character is not feasible for most encryption tasks in practice.

2.2.4 Control cf Intrinsic risk

Transmission of encrypted data is a partial telling of the secret key. It narrows the list of possible keys, thus making it easier to make an educated guess. This partial telling is a control risk, and the sender has an on-going choice of whether to transmit data or not. The guessing itself is the intrinsic risk - the sender has no influence over how much time or effort a rogue expends in guessing.

This theoretical distinction blurs the responsibility for a rogue guessing the secret key, because it has both control and intrinsic elements. Fortunately, in practice the control risk of outright theft is generally so much greater than that of cracking the code that the increased risks due to transmission may be ignored.

However, this marginal influence over the intrinsic risk does suggest that the risk of guessing should be allocated to the sender - or whoever determines the amount of data to be sent. This is analogous to the party choosing the secret (and thus also choosing the intrinsic risk) bearing the risk of guessing\textsuperscript{14}.

Achilles heel

As Mark Sneddon notes (a member of the Expert Group on Electronic Commerce)\textsuperscript{15},

"although the logical security of digital signatures based on public/private keys is very high, the weakest point in the authentication security chain is not the cryptographic security of the private key but the physical and logical security of the access device through which the private key is operated."

\textsuperscript{14}The insight that the choice of identification mechanism was significant was given me in a personal communication with Michael Attipa, and then reinforced by the test of whether a 'security procedure' was 'commercially reasonable', as stated in: Deborah L. Wilkerson, Electronic commerce under the U.C.C. section 2-201 statute of frauds: are electronic messages enforceable? (1993) 41 University of Kansas Law Review n2 p403-429, p. 422

CHAPTER 2. AUTHENTICATION TECHNOLOGY

For instance, a long private key stored on a smart card has a very high intrinsic security - but the same control risk as any physical card. If the card is accessed by four-digit PIN, it has the same control risk as an EFT debit card.

The control risk is the Achilles heel of encryption, and is often overlooked.
2.3 Addresses

If a single telegraph wire connects you and me, any communication that you receive delivered via that wire must be from me; provided that I retain control of my end, and a rogue does not divert the wire and introduce false messages.

2.3.1 Callback

A telephone number can also authenticate as an address, though the mere fact of receiving a telephone call tells you little about the sender - it could be from almost anyone. But you can authenticate them to some extent by calling them back; their telephone number then acts to identify them.

The internet addresses of web sites and email can also authenticate: your browser sends a request to the http address of a web-site, and the web-site responds by sending you the information comprising the site. The response of the web-site to your request is a form of callback. The control risk is that a rogue takes possession of the web site’s http address; the intrinsic risk is that the request for the web-site is intercepted, and a fake web-site is sent in reply. In this example, the origin of the http web-site is authenticated to the recipient, who also initially requested the web site.

In contrast, three messages are required when a sender initiates the sending. Consider authentication by delivery to an email address: I send you email containing the message; you reply to my address, requesting confirmation; I affirm it.

The control risk for email is that I could lose control of my email account, by a rogue learning my login name and password, for example. After sending an initial message apparently from me, this would enable the rogue to receive the request for confirmation; and so to reply to it.

Another scenario is that a rogue intercepts the request for confirmation before it reaches me. This interception and substitution - or diversion - is the intrinsic risk of authentication by address.

2.3.2 Internet Addressing

Effective callback requires reliable delivery. However, delivery to internet addresses is less reliable than postal or telephonic delivery because it is open and distributed, making interception much easier.

Web sites and email use the same two-tier addressing system: Domain Names and Internet Protocol addresses. The Domain Name (DN) is the user-friendly version, such as:

- http://www.law.monash.edu.au

while the Internet Protocol (IP) address is the address actually used by the internet to deliver requests. For example, the Internet Protocol version of the above Domain Name is:
http://130.194.11.21

Delivery to an Internet Protocol address bears a specious similarity to postal mail, in that the request passes through various centres, which choose the appropriate route for delivery. Unlike postal mail, these centers need not be hierarchical, so that there may be many possible routes connecting any two addresses on the internet, passing through many different nodes. Messages can be diverted at any one of these nodes.

However, nodes do tend to form a hierarchy, and so different nodes conduct different amounts of traffic, depending both on the popularity of the routes they form, and on their level in the hierarchy. Centres or routers with large “catchment” areas are generally made more more secure than those carrying less traffic, because an attack would be more disruptive.

**Domain Name Servers**

The translation from Domain Name to Internet Protocol address is provided by Domain Name Servers. By subverting a Domain Name Server, a rogue can divert a request for a web site to a different Internet Protocol address, by providing that address instead of the correct translation. Thus, the user-friendly Domain Name layer creates a second level of intrinsic risk: the rogue could divert delivery of the actual IP address; or cause diversion by giving a false translation from DN to IP address.

The Domain Name Servers are distributed throughout the internet. While this prevents the entire network from being subverted in one fell swoop, it presents many, many points of vulnerability, in the form of the many Domain Name Servers, nodes and routers.

**2.3.3 Forging Return Addresses**

Consider: a message delivered to you cannot in itself authenticate its originating address. If the message contain lies about its sender’s identity, it may lie about its originating address, too. Although email includes a field showing the address of the “sender”, this is really only the “purported sender”\(^\text{16}\).

As with web-sites, there are two levels of return addresses that may be forged: the Domain Name or the Internet Protocol address. The Domain Name return address may be forged using the “sendmail” program on a unix machine\(^\text{17}\), for example. As

\(^{16}\)Andrew Jablo, “God mail: authentication and admissibility of electronic mail in federal courts American Criminal Law Review Summer 1997 v34 n4 p1387-1408

\(^{17}\)For example:

telnet MACHINENAME.cs.monash.edu.au 25
HELO rubbish
MAIL FROM: "fake name <fake@fake.domain.name>"
RCPT TO: recipient@their.address
the Domain Name is what most people notice, forging this will have more effect on the average user than changing the Internet Protocol address, which is usually hidden in the headers\textsuperscript{18}.

Forging Internet Protocol addresses is too complex to go into here, but it is possible.
CHAPTER 2. AUTHENTICATION TECHNOLOGY

2.4 Conclusion

One purpose of the foregoing discussion of tokens, secrets and addresses has been to illustrate the universal application of the control and intrinsic risks. These suggest a simple allocation of liability: the sender is liable for a loss control of the authentication mechanism - the token, secret or address - and the intrinsic risk is borne by the party who chooses the identification mechanism. In effect, this party insures against the intrinsic risk.

However, this suggested liability assumes that a recipient has been already been fooled by a rogue; it is based on the appearance from the recipient’s point of view, rather than the reality. For example, if a rogue impersonates me, and sends a message to you which appears perfectly authentic, and you rely on this to your detriment, the question of liability comes up. Under the conventional forgery rule for signatures, I am not liable - I did not send it. In contrast, I would be liable if I had lost control, by the above suggestions. This is recipient-based liability, which has serious problems where there is a large number of possible recipients, as the next chapter shows.
Chapter 3

Administration Risks

An identification system may administrate the identification technology in several ways: by certifying the identity of its subscribers at the time of application for the identification; by implementing loss limitation devices; and by facilitating revocation of a compromised identification mechanism.

In other words, the identification system applies in the morning, limits in the afternoon, and revokes in the evening.

3.1 Application

If an identification mechanism is to identify a actual person, rather than merely consistently identify the same person, there must be some method of ensuring that the identification of the person is accurate in the first place. That this identification is wrong is the application risk.

A phone book is an example of a statement of who a person is: the phone number is the identification mechanism, and name and address is who the actual person is. Of course, names and addresses are themselves merely identification mechanisms, and are likewise subject to control, intrinsic and application risks. Their application risk is in turn based on still other identification mechanisms.

The methods of ensuring identification include birth certificates, passports, Medicare cards, driver’s licenses, delivered letters and utility bills and the like.

However, it has been argued that commerce does not require actual identification\(^1\). Sometimes only an attribute of identity, such as creditworthiness, or that funds be available to satisfy judgment, thus facilitating legal recourse. As noted in Chapter 1, sufficient reputation and recourse for commerce may be available without actual identification. That is, once a reputation is established on-line, that itself, and protection of it, may be sufficient for commerce. Since availability of such legal recourse would require evidence of the transaction, this would make it amenable to

\(^1\)Though the ATO Report focuses very much on this application aspect of identification.
regulation, including taxation.  

3.2 Limitation

Loss limitation devices seek to limit the damage caused by an impersonator. It is important to limit this reliance, both to prevent damage and to discourage impersonation, for no matter how strong our security, we must acknowledge the possibility of breach, and manage its consequences.

An individual party can easily limit their own reliance by refusing to rely beyond a certain limit. Other limitations are possible for the relying party to enforce, such as fixed duration of validity, such as, the expiry date of credit cards. The authentication mechanisms might be limited to a particular type of transaction, or geographical or other location. On problem with this approach is how the limitation is communicated to the relying party. If the identification mechanism itself states the limitation, then a rogue who can forge the identification may also forge the limitation information. This consequence of the intrinsic risk should be taken into account when selecting the security of the mechanism. However, there is a much more fundamental problem which limiting by the individual party cannot address.

Where a person uses one mechanism to identify themselves to a range of others, the effect of compromise of this mechanism is multiplied. Each individual recipient can limit their own reliance, but the total damage is all their reliance combined; a rogue can benefit greatly by posing as such a person.

In a one-to-one relationship, as between the sender and recipient of an electronic message pursuant to an EDI agreement, the recipient can limit the consequences of damage. For example, when a customer sends instructions to his bank, the bank needs to identify the customer before executing the instruction. For such instructions in the form of a cheque, the identification mechanisms are the anti-counterfeit measures in the chequebook, and the signature of the customer. Because the bank is the only recipient of these instructions, it is a simple matter for the bank to limit its reliance; such as to the customer’s account balance.

But in a one-to-many relationship where many parties rely on one method of identification, such as between a credit card holder and a merchant, the consequences are multiplied. For example, the old-fashioned “click-clack” credit card imprint machines usually required no authorisation up to a floor limit of $50. Thus, there

\[2\text{There is also the possibility of offshore legal recourse may encourage forum shopping, and thus expose regulatory schemes to competition.}\]

\[3\text{The quantification of reliance where loss has crystallised would be a question for the court; recipients need only ensure that their reliance is less than their limit.}\]

\[4\text{This danger is partially adverted to in the ECEG Issues Paper No.1, in Part 2(B)(2)(ii)(e)(i):

"that a CA would not necessarily know the uses to which a certificate might be put, in terms of numbers and value of transactions (and therefore potential size of possible liability), third-party liability insurance at a reasonable cost may not be available."}\]
was no identification of the customer by the credit card company at the time of the
transaction. This allowed a rogue to enter many fraudulent transactions, possibly
many times over the credit limit of the card.

Fortunately, the total reliance in such a one-to-many relation can be limited: by
limiting the number of relying parties; by each relying party using a different and
independent identification mechanism; and by the relying parties communicating
their reliance amongst themselves. If the maximum number of relying parties is
limited to a relatively small number, then the total reliance is limited to their combined
reliance - a relatively small multiple. For two recipients, the total reliance is double
the individual limit. For large numbers of relying parties (consider twenty million),
or where the number is indeterminate, this limitation is practically useless.

If each party uses an identification mechanism independent of the other to iden-
tify the sender of a message, then the consequences of compromise of just one of
these mechanism is again limited.

If the two parties communicate, they may sum their individual reliances to deter-
mine whether the agreed limit has been exceeded. A convenient way to accomplish
this is via a common point: each act of reliance is communicated to this common
point, which keeps track of the total reliance. When a total (or daily) limit is
exceeded, later parties seeking reliance are advised of the danger.

3.2.1 Breadth

The breadth of the internet presents virtually no limit to the number of recipients
that an imposter could deceive. In a world of global transactions, instant and
automated, the aggregate consequences of impersonation may be very large indeed.
Although an ideal system of identification would be good to the world at large, one
loss limitation solution is to restrict internet commerce to one-to-one relations, or
one-to-many relations where the number of recipients is small.

Though this is a problem of the internet environment and not of the actual
identification mechanisms used (such as electronic signatures), the issue of breadth
is also raised in traditional commercial environments.

In person The number of people who know a person by sight is usually limited
to something of the order of one thousand. While this is not a small number, it is
much smaller than the number of potential recipients on the internet. Additionally,
impersonation is much more difficult in person than on the internet.

Credit card transactions in person Service stations on busy roads do not
expect to know their customers by sight; and so a single rogue could travel to
many such outlets pretending to be someone else. Where no authorisation is sought
(as when the transaction is less than $50, or the authorisation mechanism is out of
order), this fraud is limited by geographical distance and time spent moving between
merchants. The risk of being caught out in some way would also mount with each
transaction. Clearly, a rogue can only travel to so many petrol stations before the theft is reported; each visit is a risk of being discovered; and each visit leaves a witness behind who may recognize him. These factors prevent large aggregate losses being accumulated through transactions under the $50 limit.

**Credit card transactions in distance selling** In contrast, in distance-based transactions such as mail order, phone order, and paying for long distance telephone calls, there is little alternative identification apart from the credit card number; there is little chance of being caught; and less time is needed between transactions. Today, all such transactions are verified before proceeding, so the inevitable fraud is limited.

The internet is the nightmare version of these risks. However, one solution is to use many identification mechanisms, each limited in breadth, for example, by geography. A danger with this method is that the part of the mechanism that states its geographical validity (or type of transaction or other restriction in breadth) may itself be forged. But these apparently independent mechanisms have another risk.

### 3.2.2 Independence

Using more than one identification mechanism - perhaps a different one for every recipient, or a different one for each class of recipient (perhaps based on geographic location) - can reduce the reliance on each one. But this solution is not as effective as it might seem.

Although the problem is stated as where “a person uses one mechanism to identify themselves to a range of others”, the problem is not exactly one of using the *same* identification mechanism; it is the risk that the means used to identify by several people is compromised. For example, if I have 10 different credit cards in my wallet, losing my wallet loses them all. They each identify me to different people; but if one is compromised by theft, they probably all are. Having all your eggs in one basket in this sense is similar to storing your EFTPOS card and PIN together, so they are susceptible to simultaneous theft. That is, that the event allowing compromise of the one also facilitates compromise of the other. Their control risk is not independent.

Analogous to losing my wallet is storing all encryption keys on the same hard drive on a computer, so that a hacker breaking in will gain access to them all.

The intrinsic risk may also be dependent. Authentication mechanisms may overlap by resting on the same underlying identification mechanism: a mathematical flaw being discovered in the encryption algorithm used by all the identification mechanisms would simultaneously expose them all to increased intrinsic risk. Compromise of an email address (either by diversion or by taking control of it) would compromise all mechanisms insofar as they use this address: call-back for one person will generally be to the same address.

The point is that identification mechanisms should be independent so that loss of one does not entail loss of the other. Although the risk for different mechanisms
can be separated to some extent (for example the \textit{intrinsic} risk of encryption keys is not increased by storing them together), some overlap is inevitable. Having more than one identification mechanism is not a perfect solution to the loss limitation problem.

3.2.2.3 Registering Reliance

Reliance on an identification mechanism can be limited by use of a common point which keeps track of the total reliance by each reliance being registered there - but this common point must itself be identified. For example, a merchant first identifies the customer with a credit card as being the cardholder through possession of the card and by their signature; but before relying on this identification, the merchant contacts the credit card company for authorisation. The credit card company acts as a common point, by tracking the aggregate reliance on the card, and limiting this reliance to a total or daily limit. This common point must in turn be identified: the merchant must be able to identify the credit card company's authorisation of payment as originating from the credit card company.

The danger is that the credit card company could be impersonated when the merchant seeks verification. Because many merchants rely on this one point to verify their many customers, reliance is concentrated. In contrast, an individual credit card is used to identify just one customer.

The identification of the credit card company may be analysed in a similar way to identification of the customer: a large number of relying parties (merchants) multiples the consequences of compromise. A merchant telephoning the company for verification could be diverted to the rogue. A band of accomplices could present several stolen credit cards at a number of stores, confident that they would be verified. Needless to say, this is a complex and organised fraud - but in the internet scenario, one rogue could perform all roles, and automate much of it.

The same loss limitation devices already discussed can be considered in relation to this identification of the credit card company.

\textbf{Breadth} The credit card company could limit the number of merchants relying on identification of it.

\textbf{Independence} Each merchant could use a different identification mechanism to authenticate the credit card company. For example, a different phone number; though this has the weaknesses already discussed.

\textbf{Registering Reliance} Unfortunately, the credit card company already is a common point for the relying merchants; and so this cannot be used again.

But the simplest response to greater reliance (through the many merchants' reliance on many customers) is greater security: the identification of the common
point might employ a more secure channel of communication than the internet, such as connection via a telephone line or an independent network.

Limiting identification reliance by registering that each reliance at a common point, which itself requires identification, is answering a riddle with the same riddle. It is not a solution. While it may be possible to go beyond the internet to provide the additional security needed (such as a telephone connection), this is no longer the self-contained electronic commerce that has been envisioned.

3.3 Revocation

Revocation is as simple as warning someone that an identification mechanism is no longer reliable: for example, cheques and credit cards may be stopped.

Where there is reliance by multiple parties, a common point can facilitate revocation. For example, a merchant seeking authorisation for a revoked credit card will be advised to not rely on it by the common point.

Generally, the originator of a notice of revocation need not be identified as stringently as the actual identification mechanism itself, as the consequences are less serious. For credit cards, knowing the credit card number and perhaps some secret password (such as your mother's maiden name) is generally sufficient. This is justified by the less serious consequences of impersonation in this context: stopping a cheque or credit card may inconvenience their owner, but they would not lose money directly, and it would not reward a rogue to engage in this impersonation.

It is important that revocation be effected as quickly as possible, to attempt to limit the damage that a rogue may do. The surest way is to require authorisation from a common point before relying on the identification mechanism.

3.4 Conclusion

An alternative means of identification is required all three phases of an identification system: limitation of aggregate losses requires identification of a common point; application requires some proof of identity before the mechanism is created; and a notice of revocation should only be accepted from the rightful holder of the identification mechanism - but generally they revoke because they have lost control of the mechanism, and therefore cannot use it to identify themselves. Limitation is by far the most important of the three.

The methods of limiting reliance by limiting audience, using independent mechanisms and common points do not seem able to fulfill the dream of internet commerce, which requires broad identification to establish trust5. However, identification within small groups does seem feasible; and this may provide a way to approach the internet dream.

5see Chapter 1
3.4. CONCLUSION

Without some form of limitation in the ways suggested, an identification mechanism may generate indeterminate reliance, and thus an indeterminate total liability. Shifting this liability to the recipients, the purported sender or the common point does not make it bearable. Recipient based liability in an open environment where total reliance is not limited simply will not work.

The task for the internet is for the security of the common point to be sufficiently high in comparison with the total possible gains to a rogue. This is a difficult commercial balancing if it is possible at all, and is best left to industry.
Chapter 4

Article 13

The approach taken by Article 13 of the UNCITRAL Model Law is partially recipient-based: sometimes, you would be held to your word, even if you did not say it.

Article 13. Attribution of data messages

(3) As between the originator and the addressee, an addressee is entitled to regard a data message as being that of the originator, and to act on that assumption, if:

(a) in order to ascertain whether the data message was that of the originator, the addressee properly applied a procedure previously agreed to by the originator for that purpose; or

(b) the data message as received by the addressee resulted from the actions of a person whose relationship with the originator or with any agent of the originator enabled that person to gain access to a method used by the originator to identify data messages as its own.

(4) Paragraph (3) does not apply:

(a) as of the time when the addressee has both received notice from the originator that the data message is not that of the originator, and had reasonable time to act accordingly; or

(b) in a case within paragraph (3)(b), at any time when the addressee knew or should have known, had it exercised reasonable care or used any agreed procedure, that the data message was not that of the originator.

4.1 The Operation of Article 13

Article 13(3)a confines the appearance that a recipient must examine to a pinpoint of identification while 13(3)b focuses on whether this appearance was created by
the sender’s loss of control. These presumptions may be rebutted by notice and knowledge, as per 13(4)a and 13(4)b. The recipient is not entitled to rely on a message if they receive notice that the message was fake, or if they knew that the message was fake. This knowledge exception does not apply to an appearance under 13(3)a.

One interpretation of Article 13 is that it attempts to simultaneously regulate identification in both closed and open environments in one fell swoop: the appearance of 13(3)a applies to closed environments, while the control of 13(3)b and knowledge exception of 13(4)b apply to open environments\(^1\). This interpretation resolves the curious overlap between appearance\(^2\) and control\(^3\), and the odd effect of knowledge\(^4\) rebutting only the control provision.

A possible reason for this combination is a wish to roll the signature and EDI models of attribution into one. When these models are kept distinct, the result is merely awkward; but where they are conflated, their goals are both frustrated.

4.1.1 Appearances

Article 13(3)a narrows the focus to a single pinpoint of identification - and ensures that the allocation of liability is partially recipient-based. A recipient is entitled to rely on an impersonation when the performance is so good that it that it fools an agreed authentication procedure. This strictly defined test is convenient for the recipient, and secure enough for the purported sender that it agreed to it.

(3) As between the originator and the addressee, an addressee is entitled to regard a data message as being that of the originator, and to act on that assumption, if:

13(3)a in order to ascertain whether the data message was that of the originator, the addressee properly applied a procedure previously agreed to by the originator for that purpose; or

Authentication Procedure

An authentication procedure is the mechanism used by the recipient to identify the sender of a message. This concentrates reliance into the appearance from the recipient’s point of view, as defined by the procedure.

Not all authentication procedures are equally secure. The ECEG note that the parties should not be completely free to agree to any authentication procedure, to protect the weak from stronger parties who would force a choice of procedure that

\(^1\)The notice exception of 13(4)a applies to both open and closed environments.
\(^2\)13(3)a
\(^3\)13(3)b
\(^4\)13(4)b
places an unfair risk on the other\textsuperscript{5}. There is a real risk of this\textsuperscript{6}. There should be 
standards of “security and reliability”\textsuperscript{7}, and of what is “fair and reasonable”, perhaps 
based on s.68A(3) TPA 1974\textsuperscript{8}. Note that if a weaker party relies on a message sent 
by a stronger party pursuant to a below-standard procedure, the stronger party may 
be able to repudiate its message on the basis that the procedure was inadequate: 
the suggestion may back-fire.

**Agreed Procedure**

Agreement on the procedure is important for a purported sender to be bound fairly. 
Such agreement necessarily includes an agreement to be bound, as that is the purpose 
of the procedure. This agreement on procedure may be directly with the recipient, or 
through a third party, such as a carrier. However, agreement appears to be also 
needed in another sense, for a different purpose: to provide a closed environment.

The appearance provision also requires messages to be “based on a previous 
agreement”, and this agreement must be between the sender and recipient: this 
provision “does not apply in an open environment” but only where there is a “pre-
vious agreement”\textsuperscript{9}. For the environment to remain closed, this second agreement 
must refer to an agreement between the originator and recipient, and not with a third 
party (such as a carrier), in contrast with the first agreement. A closed environment 
limits the number of possible relying recipients, unlike the open environment of the 
internet, which has an indeterminate number of potential recipients, as discussed in 
Chapter 3.

Unfortunately, Article 13 does not explicitly distinguish between these two agree-
ments, and the division into open and closed environments, and is here implied from 
the commentary.

### 4.1.2 Control

Article 13(3)b looks at how an appearance was created, in terms of how well the 
originator controlled its method of identification:

\textbf{13(3)b} the data message as received by the addressee resulted from the 
actions of a person whose relationship with the originator or with any 
agent of the originator enabled that person to gain access to a method 
used by the originator to identify data messages as its own.

If the originator fails to take reasonable care\textsuperscript{10} to control the method of access, the 
addressee is again entitled to regard the message as from the originator. This duty

\begin{itemize}
  \item \textsuperscript{5}4.5.66 p. 171
  \item \textsuperscript{6}4.5.79, p. 174
  \item \textsuperscript{7}4.5.66-67 p.171
  \item \textsuperscript{8}4.5.79, p. 174
  \item \textsuperscript{9}2.13.4, p. 85, from the Guide to Enactment which accompanies the UNCITRAL Model Law
  \item \textsuperscript{10}2.13.5, p. 85, Guide to Enactment
\end{itemize}
to control extends to prevent a rogue gaining access by virtue of their relationship with the originator (or their agent). The onus is on the alleging party.\(^{11}\)

Note that this provision does not exactly coincide with the definition of “control risk” used in this thesis. Consider a rogue who gained access to an identification method and impersonated the originator, but who had no relationship with the originator or its agents. This would constitute breach of a control risk; but would not fall within Article 13(3)b with the consequence that originator would not be bound, despite being in the best position to prevent this breach.

**Appearance of Control**

The *appearance*\(^{12}\) and *control*\(^{13}\) provisions differ in scope: the *appearance* provision seems to apply to closed environments\(^{14}\) (where the communication is based on a previous agreement and where the originator has agreed to an authentication procedure), whereas the *control* provision applies generally, including to open environments (where the communication is not based on a previous agreement or where the originator has not agreed to an authentication procedure).

Thus, in an open environment, the originator is bound only through failure of his control. This corresponds to the rules for handwritten signatures: the apparent signer is not liable for forgeries; but is liable where he lost control of the identification device. In the case of handwritten signatures, it is impossible to lose control of one’s hand; except in the sense of one’s will being overborne or influenced.

One comparison between these provisions is that the appearance provision is based on *appearances* from the recipient’s point of view, whereas the control provision looks to the reason for that appearance, in terms of the sender’s loss of control. However, if the loss of control facilitates the appearance of authenticity, the appearance provision operates, and the control provision is unnecessary. That is, if the appearance of authenticity is not created, so that the appearance provision does not apply, then the recipient should not have been relying on the message in the first place. Should this foolish recipient be saved by the purported sender happening to have lost control? Similarly, a foolish addressee who did not apply the procedure properly (and thus is not covered by the appearance provision) may be saved if the originator fortuitously turned out to be careless by losing control of its method of access.

The interaction between the appearance and control provision is thus unclear, for a failure to safeguard a method of identification would facilitate a rogue sending a message satisfying an agreed authentication procedure, making the originator liable under the appearance provision. That is, the two provisions seem to have some overlap; but it is unnecessary to have both if they cover exactly the same situation.

\(^{11}\)The difficulty and expense of proving such negligence is not discussed.

\(^{12}\)13(3)a

\(^{13}\)13(3)b

\(^{14}\)2.13.4, p. 85, from the Guide to Enactment
4.1. THE OPERATION OF ARTICLE 13

The suggested difference in scope resolves this curious interaction: the appearance provision applies to closed environments, while the control provision applies in open environments. In an open environment, a failure of the apparent sender’s control of his means of identification is the only impersonation that may bind him. The recipient bears the intrinsic risk; and this should motivate him to gather more information to assure him of the identity of the originator of the message. There is of course the further problem of how the recipient may prove that it was a loss of control and not an intrinsic risk that facilitated the impersonation.

4.1.3 The Notice exception

The presumptions raised by the above provision may be rebutted by notice and knowledge, per 13(4)a and 13(4)b.

13(4) Paragraph (3) does not apply:

13(4)a as of the time when the addressee has both received notice from the originator that the data message is not that of the originator, and had reasonable time to act accordingly; or

The recipient is not entitled to rely on a message if notice is received that the message is fake. Notice changes the appearance from the point of view of the recipient, and so is consistent with the appearance provision.

4.1.4 The Knowledge exception

13(4)b in a case within paragraph (3)(b), at any time when the addressee knew or should have known, had it exercised reasonable care or used any agreed procedure, that the data message was not that of the originator.

The recipient is not entitled to rely if they knew that the message was an impersonation. This knowledge exception has three limbs:

1. if the addressee knew that the message was not sent by the originator;
2. if the addressee should have so known, had it taken reasonable care; or
3. if the addressee should have so known, had it used an agreed procedure.

Not applicable to appearance

This this knowledge exception does not apply to the appearance provision. That this could sanction “unconscionable conduct or at least wilful blindness”\(^\text{15}\) is not acceptable to the ECEG. That is, an addressee might know that a message is fake,

\(^{15}\)ECEG p. 172
but still be entitled to rely upon it, if it satisfies the procedure; or discourage an addressee from taking the reasonable care that that would yield this knowledge.

UNCITRAL sees preserving the reliability of procedures as worth this risk. Perhaps the concern is the expense of taking reasonable care and maintaining proof of this and discouraging disputes. That is, the certainty and efficiency of a bright line confining disputes is worth some injustice. Note that the appearance provision explicitly requires pre-existing agreements, and the parties should be expected to make their own bargain here. The protection of reasonable care is only available where it will not upset the parties’ own agreement: it is a default.

In contrast, the ECEG views this injustice as too high a price to pay - a party should not be held to “their” word if they did not say it, unless the relying party took reasonable care to ensure that it was their word; a fortiori if the relying party knew it was fake. However, the question arises that in practice, in an international internet transaction, what other indicia of identity would be “reasonably” available? The whole purpose of the scheme is to enable reliance on a single pinpoint of identification.

Reasonable care The appearance provision concentrates risk into the authentication procedure, and so broadening the appearance that the recipient must consider would frustrate this purpose. Therefore, the “reasonable care” limb should not apply to the appearance provision.

The reasonable care limb of Article 13(4) dilutes reliance on the pinpoint of the authentication procedure - the recipient is not protected by relying on the authentication procedure alone. This exception discourages complete reliance on any identification method or procedure, as reasonable care must also be exercised. It prevents all the eggs being in the one basket, and softens reliance on the method used by the originator to identify data messages as its own.

Though “reasonable” is a flexible concept, any such compromise is destructive, for in shifting reliance away from the single pinpoint, it destroys the convenience of a strictly defined identification; and yet also fails to safeguard the innocent sender.

Agreed procedure The “agreed procedure” limb duplicates the requirement already present in the appearance provision, so it does not matter whether this limb applies to that provision or not.

Actual knowledge However, the first limb, of actual knowledge, would not dilute the pinpoint, but prevent out-and-out fraud.

---

16. 13.7 Guide to Enactment
17. Article 13(3)a
18. 3(b)
4.2. THE ECEG POSITION ON ARTICLE 13

Recommendation

The limb of actual knowledge could be simply severed, so actual knowledge is an exception to both the appearance and control provisions\(^{19}\), leaving the other two limbs to apply only to the control provision.

4.1.5 Consequences of Attribution

13(3) As between the originator and the addressee, an addressee is entitled to regard a data message as being that of the originator, and to act on that assumption...

To what extent is the recipient entitled to rely on the word of the purported sender? The Guide to Enactment \(^{20}\) says that this article does not assign “responsibility”, but merely establishes a presumption that a message would be considered as of the originator, the ECEG notes that the Article does not “directly assign the consequences of responsibility” \(^{21}\).

The consequences of attribution may be agreed on beforehand by the parties. A pre-existing agreement may specify the effect of being entitled to rely, or not being so entitled. Indeed, some kind of agreement, both for the procedure, and the message, is required by the appearance provision; though not for the control provision, which may operate in open environments. But default rules are required to cover situations not contemplated by these agreements; or where there is no agreement.

4.2 The ECEG position on Article 13

Should you be held to your word if you did not say it?

The position of the ECEG is that the sender’s point of view wins over that of the recipient. That is, you will only be bound if the appearance that it is your word meets the high standard of fooling a court of law, given all the evidence available. The ECEG rejects both the Article 13(3) and 13(4) that have been discussed.

The Argument from Tradition

The conservative argument objects to electronic commerce being treated differently from traditional paper-based commerce without good reason\(^{22}\). The existing system has existed for centuries, and seems to work pretty well - it the devil we know. A new system, with unknown problems and consequences should be approached with the respect and caution deserved of a Pandora’s Box.

\(^{19}\) The agreed procedure could also apply to both provisions, but this would be redundant, because the appearance provision explicitly involves such a procedure.

\(^{20}\)2.13.1, p. 84, paras 83-92, pp 46-48

\(^{21}\)Footnote 165, p. 84

\(^{22}\)4.5.77, ECEG, p.173
In paper-based commerce at a distance, there is a risk of impersonation, usually in the form of a forged or unauthorised signature. In general, the addressee can rely only on the reality - if the signature was forged, it does not matter how good the forgery was, the apparent sender is not bound.

The law of agency provides an exception to this, where a genuine signature affixed without authority allows the purported originator to be bound by the appearance of authority.

The question of whether the apparent signer is bound is resolved by whether he signed or not. Both parties are free to adduce evidence for this determination\(^{23}\).

**The Argument from Efficiency**

The ECEG notes that a legislative allocation of commercial risk may involve “pre-emptive assumptions”\(^{24}\) about fair and efficient business practice, which may well be wrong. Secondly, different types of commerce and different types of authentication methods are likely to require different allocations of risk, for efficiency and fairness. Any broad legislative allocation across a wide commercial context may not be right for them all\(^{25}\). A sweeping and pre-emptive allocation may have “serious unintended consequences”\(^ {26}\).

**The ECEG’s conclusion**

The ECEG suggests specifically that a purported sender can be held to their word only when they are proven to have sent it. This places the risk on addressees, who should then require “reliable authentication methods” or seek “additional authentication indicia which create a strong evidential basis that the apparent originator did send the data message”. The onus is on the addressee\(^ {27}\).

In effect, this shifts the problem to the courts to evaluate the probative weight of the various indicia of authorisation. This thesis is relevant to the exercise of that judgement.

Very importantly, the ECEG does not oppose attribution rules *per se*, but only to such rules imposed on parties by legislation. Parties are free to agree to attribution rules by contract\(^ {28}\), for example by a trading partner agreement\(^ {29}\).

However, reliance on this agreement is disallowed, unless it is “fair and reasonable to do so in all the circumstances”\(^ {30}\). This protects weaker parties from having unfair agreements imposed on them by stronger parties.

\(^{23}\)4.5.78, ECEG, p. 173
\(^{24}\)4.5.76, ECEG, p. 173
\(^{25}\)4.5.76, p. 173, and 4.5.79, p.174, ECEG
\(^{26}\)4.5.76, p.173
\(^{27}\)Recommendation 12, ECEG p.175
\(^{28}\)Recommendation 12, ECEG, p. 175
\(^{29}\)4.5.78, ECEG p.174
\(^{30}\)Recommendation 12, ECEG p. 175; see also earlier discussion under Article 13 of the UNCI-TRAL Model Code
4.3 Conclusion

The ECEG is right to revert to the signature model of traditional paper-based commerce, in an open environment. Equally, it recognises that in a closed environment - where there are pre-existing agreements - parties should be free to allocate risks themselves. The ECEG recognises the UNCITRAL’s folly of attempting to combine the signature and EDI models to solve the riddle of identification on the internet. Instead, the ECEG insists that these models must be kept distinct for their proper operation.

Unfortunately, the riddle remains that the UNCITRAL sought to solve: how may a recipient rely on its identification of the sender of a message received on the internet?

It is the thesis of this dissertation that the correct approach to solving this riddle is the credit card model; that is, a common point is the only way to limit the liability for attribution of messages on the internet. This solution relates to the UNCITRAL Model Law and the ECEG report as a component in determining whether an authentication procedure is “fair and reasonable”. Certainly, indeterminate liability is neither fair nor reasonable.
Chapter 5

Conclusion

The concepts of sender and recipient-based liability, and control and intrinsic risks are brought together to construct a principled basis for legal rules for allocating risks of identification in commercial transactions on the internet.

Sender or recipient-based?

Sender-based or recipient-based liability is a matter of choice; but if recipient-based liability is chosen in an open environment, reliance registration is a matter of necessity.

Once a system of reliance registration is assumed, a more intricate allocation of liability is possible. Examples of such systems include EFTPOS cards, cheques, credit cards and the Torrens Land System.\(^1\)

Control The blunt allocation is to assign losses to the party best able to prevent them, thus motivating them to prevent the loss. Therefore, the party having the control risk then bears the losses caused by their loss of control. The allocation of loss flowing from intrinsic risks is more flexible, as neither party can prevent them.

Influence However, a party may have influence over the intrinsic risk. One suggestion is to allocate the intrinsic to the party with influence over the size of the risk through the choice of identification mechanism, thus motivating them to choose one with minimum intrinsic risk. But who chose it? Freedom of contract says that the parties agree - they both chose it. In practice, unequal bargaining power suggests that the stronger party would select it; i.e. the provider, not the consumer. An alternative is to allocate it to the party who can most efficiently obtain insurance, which yields the same conclusion.\(^2\)

---

\(^{1}\) The closed environment of EDI may be thought of as registration, since the one recipient is able to register its own reliance.

\(^{2}\) A fine distinction arises: for example, each use of an encryption key presents a control risk, because knowing the encrypted data can help infer the secret key - using a key partially reveals it.
CHAPTER 5. CONCLUSION

Stipulating a minimum standard of security explicitly removes the choice to a large extent, leaving only the party who can most efficiently insure to bear the loss. Generally, this will be the provider, who is usually also the stronger party.

Consider a third party providing a reliance registration system using 1024 bit asymmetric encryption keys. The blunt approach allocates the risk of loss of key to the holder; whereas the the risk of a rogue inferring the key is to be born by the provider, who selected the 1024 bit length. Thus, holders are motivated to safeguard their key; and the provider is encouraged to use longer keys, which are more difficult to infer.

A smart card containing the key and accessed by a 4 or 6 digit PIN has the control risks of loss of the PIN; loss of the card; and loss of the key - and the intrinsic risks of a rogue guessing the PIN; counterfeiting the card; and inferring the key. There are also more subtle risks: power usage and electromagnetic emanations from the card itself can be used to help a rogue determine the key. Note that emanations from a smart card are a general security issue, and not specifically one of authentication. That is, the keeping of a secret key is an identification mechanism; but the means of keeping it secret is a security device, which does not identify in itself; it plays merely a supporting role.

Evidence

Who should be liable when the cause of loss cannot be proven? This is a default choice where insufficient evidence is available. Consider a rogue who has inferred the key contained in a smart card. The consumer can prove that the card was not lost or stolen by simply producing the card, and therefore that the key must have been inferred.

If a consumer denies transactions made with a card, how can it be proved whether the loss was due to a rogue attacking the intrinsic risk; or that the consumer is lying? Production of a card proves that the card was not stolen or lost; but does not address the above quandary.

Placing the onus of proof on the provider is in line with “innocent until proven guilty”, and motivates the provider to put procedures in place to obtain such proof. This problem already exists with credit cards: the liability up to the $50 cap motivates consumers to not practice this type of fraud. However, a consumer conducting much larger transactions could still gain; but would quickly attract the attention of the provider if it was repeated.

The problem of evidence is a complex one, which requires more study in the context of specific identification mechanisms and administration systems.

---

If this data significantly helps a rogue to infer the key, then the risk of loss is partly control and partly intrinsic. Although this transmission risk is technically a control risk, the simplest approach is to treat it as an intrinsic risk, because it must be used.
Legal limitation devices

These devices do not prevent actual losses from occurring, but merely allocate them once they have fallen. Without registration, the losses would be indeterminate, and unbearable and uninsurable by any party, and allocating such liability would be merely an exercise in rearranging deckchairs.

Fortunately, registration of reliance does solve this problem, as it has done for credit cards and EFT cards; and the legal limitation devices employed there may be borrowed.

Standard of Care  The standard required by the above blunt approach is strict liability - the holder is liable for any loss of control. Alternatives are reasonable care and the higher standard of diligence. Of course, these are merely imprecise labels, and a specific identification system is required in order to determine the standard precisely. However, where consumers are involved, a standard of reasonable care seems the most reasonable starting point.

Capping Liability  The liability on the holder is capped by the loss limitation devices such as a total or daily limit. These caps are only possible because reliance is registered. However, they also allow more precise allocations of liability.

Caps on liability apart from daily and total limits are also possible; for example, the $50 limit on liability for stolen credit cards. Note that this figure is quite arbitrary and only notionally related to the $50 ceiling.

These fine tuning devices of standard of care and liability caps are possible by the provider obtaining insurance, and upping fees to cover this. Whether the other, more careful consumers should be made to pay for the carelessness of a few is beyond the scope of this dissertation.

Conclusion

The three themes of this thesis are:

1. Recipient-based liability entails registration of reliance.

2. An analysis of control and intrinsic risks provides a principled and technology neutral approach to allocating liability.

3. Strong identification of an actual person (with its big brother and privacy concerns) is not necessary for commerce - weak identification is sufficient.

The logistics of a common register are a practical problem, as is the balancing of security with the consequences of failure. This thesis suggest that these may be too difficult to overcome for the large numbers of recipients in a global marketplace. However, niche markets with sufficiently small numbers of recipients are presently
plausible; and it may be possible to link them up to form larger networks of recipients.

Without a registration system, only sender-based liability is workable on the internet; but this does not provide certainty for recipients. Without this certainty, the dream of commerce on the internet will remain forever an elusive fantasy.