Carnegie Mellon University

PRIVACY & RELIABILITY IN INTERNET COMMERCE

A Dissertation
Submitted to the Graduate School
In Partial Fulfillment of the Requirements

for the degree
Doctor of Philosophy
in
Engineering & Public Policy

by

Linda Jean Camp

Pittsburgh, Pennsylvania
August, 1996

to
Shaun McDermott
strong, gentle
loving, laughing, lasting
heart of music
beloved.

Acknowledgments

I must first acknowledge Adonica Marie, without whom this dissertation may have come more quickly, but would have been written with less joy and wisdom. Wilson, who taught me many lessons, I will not forget.

My committee members each deserve individual acknowledgment. I am deeply indebted to Granger Morgan for following his own dreams and beginning the department where I have come to professional maturity. Pam Samuelson provided irreplaceable insight into the subtleties of the law, and despite her almost frightening schedule, always found time to provide detailed comments. Mary Shaw has offered valuable time and insights from her technical and personal wisdom. Bennet Yee has given both professional counsel and patient consideration. Two professors were my advisors and shared the chair of my committee, and thus deserve special recognition. This dissertation would not have come to fruition without the guidance and support of Professors Marvin Sirbu and Doug Tygar.

I am grateful to the United States Postal Service for funding. I thank Mike Harkavy for his technical insight.

To the members of my incoming class who shared my tribulations. Laura Painton and Tse-Sung Wu stand here as first among equals. Rosy Chen shared her heart, wisdom and her office. Senior student Milind Kandlikar soothed our fretful first year minds with occasional doses of perspective. Indira Nair's spirit has shaped and filled the department, and her door is always open. Donna Riley shared her rare gifts of strength and kindness, bestowed with a discerning wit, which cuts through hypocrisies both intellectual and spiritual. Ian Simpson was a source of continued intellectual excitement as his dissertation begins. Outside EPP, Richard Field offered his very relevant expertise and the kindness of his heart in reviewing and commenting on my work. Cathleen McGrath offered engaging debate or empathy, as appropriate, over uncounted cups of tea. Phoebe Sengers reminded me to like myself, and hold my work just dear enough.

Denise Murrin-Macey, Patricia Steranchak, Janice Trygar, and Victoria Massimino made my time here easier in a many ways, the greatest of which has been in the sharing of their company and friendship.

To others who have helped me, know that the absence of your name does not mean that I have forgotten your friendship. Thank you.

Contents

List of Figures v
List of Tables vi
Abstract vii

Chapter 1: Introduction & Motivation
1.1 Money, Its Functions, and Electronic Commerce 3
1.2 The Internet: The Medium of Electronic Commerce 5
1.3 Business Advantages of Internet Commerce 7
1.4 The Organization of this Dissertation 9

Chapter 2: Security & Reliability
2.1 Security 14
 2.1.1 Threats to Electronic Information Systems 14
 2.1.2 Basic Cryptographic Tools 16
 2.1.3 Availability 17
 2.1.4 Authentication 18
 2.1.5 Confidentiality 19
 2.1.6 Integrity 19
 2.1.7 Nonrepudiation 20
2.2 Reliability 20
 2.2.1 ACID Properties 20
 2.2.2 Degrees of Atomicity 21
2.3 Commerce Transactions 22
 2.3.1 Stages of a Transaction 22
 2.3.2 Browsing Information 23
2.4 Secure Hardware 25
 2.4.1 Secure Servers 26
 2.4.2 Smart Cards 27
2.5 Key Management 28
 2.5.1 Symmetric Key Management 28
 2.5.2 Asymmetric Key Management 29
2.6 Microdata Security 32
2.7 Psuedonymity & Anonymity 33

Chapter 3: Internet Commerce, Privacy & the Law
3.1 Privacy 36
 3.1.1 Codes of Ethics 38
 3.1.2 State Law 39
 3.1.3 Federal Law 41
  3.1.3.1 Statutory Law 41
  3.1.3.2 Constitutional Law 43
 3.1.4 Privacy & Information Technology 46
3.2 Cryptography Policy 47
3.3 Data Reporting & Disclosure 49
 3.3.1 Required Information Reporting 49
  3.3.1.1 Techniques for Regulatory Information Requirements 50
  3.3.1.2 Motivation of Regulatory Information Requirements 52
 3.3.2 Reporting Examples 52
  3.3.2.1 Immediate Reporting 52
  3.3.2.3 Periodic Reporting 53
  3.3.2.3 Periodic Aggregate Reporting 54

3.3.2.4 Data Storage 54
3.3.3 Reconsidering Requirements 54
3.3.3.1 Immediate Reporting 55
3.3.3.2 Periodic Reporting 56
3.3.3.3 Periodic Aggregate Reporting 56
3.3.3.4 Data Storage 56
3.3.4 Summary 57

Chapter 4: Separation & Examination of Internet Commerce Systems
4.1 Separation 60
4.2 Analysis 61
4.2.1 Transactional Reliability 61
4.2.2 Security 62
4.2.3 Privacy 62
4.2.4 Regulatory Issues 64

Chapter 5: Notational Currency
5.1 Credit Cards 66
5.1.1 A Transaction 67
5.1.2 Security 69
5.1.3 Privacy 69
5.1.4 Regulatory Issues 70
5.2 First Virtual 71
5.2.1 A Transaction 72
5.2.2 Security 74
5.2.3 Privacy 74
5.2.4 Regulatory Issues 75
5.3 Secure Sockets Layer 75
5.3.1 A Transaction 76
5.3.2 Security 77
5.3.3 Privacy 77
5.3.4 Regulatory Issues 78
5.4 Secure Electronic Transactions 78
5.4.1 A Transaction 79
5.4.2 Security 85
5.4.3 Privacy 86
5.4.4 Regulatory Issues 86
5.5 NetBill 87
5.5.1 A Transaction 88
5.5.2 Security 90
5.5.3 Privacy 90
5.5.4 Regulatory Issues 91
5.6 Anonymous Credit Cards 92
5.6.1 Transaction 92
5.6.2 Security 96
5.6.3 Privacy 96
5.6.4 Regulatory Issues 97
5.7 Summary 98

Chapter 6: Token Currency
6.1 Legal Tender: Cash 103
6.1.1 A Transaction 103
6.1.2 Security 104
6.1.3 Privacy 105

6.1.4 Regulatory Issues 105
6.2 Digicash 106
6.2.1 A Transaction 106
6.2.2 Security 107
6.2.3 Privacy 108
6.2.4 Regulatory Issues 108
6.3 Digicash with Detectable Double Spending 109
6.3.1 A Transaction 110
6.3.2 Security 111
6.3.3 Privacy 112
6.3.4 Regulatory Issues 112
6.4 MicroMint 113
6.4.1 A Transaction 114
6.4.2 Security 115
6.4.3 Privacy 116
6.4.4 Regulatory Issues 117
6.5 Summary 118

Chapter 7: An Anonymous Atomic Layer
7.1 A Transaction: Preparation & Purchase 121
7.2 Security 125
7.3 Privacy 125
7.4 Regulatory Issues 126
7.5 Further Questions 126

Chapter 8: Implications & Conclusions
8.1 Law Enforcement 131
8.2 The Business Community 132
8.3 Civil Libertarians 133
8.4 In Closing 134

Chapter 9: Bibliography 137

List of Figures

Figure 1.1
Exponential Growth of the Number of Computers Connected to the Internet 6
Figure 2.1
Cost Distribution in a Credit Card Transaction 9
Figure 2.2
Analog Equivalents of Cryptographic Capabilities 16
Figure 5.1
A Credit Card Transaction... 68
Figure 5.2
A First Virtual Transaction 72
Figure 5.3
A Secure Sockets Layer Transaction.. 76
Figure 5.4
A Secure Electronic Transaction... 80
Figure 5.5
SET with Certified Delivery... 83
Figure 5.6
A NetBill Transaction.. 89
Figure 5.7
An Anonymous Credit Card Transaction.. 93
Figure 6.1
A Digicash Transaction 106
Figure 6.2
A Digicash Transaction with Double-Spending Identifiers.......................... 110
Figure 6.3
A MicroMint Transaction.. 114
Figure 7.1
Preparation for an Anonymous Atomic Transaction 121
Figure 7.2
An Atomic Anonymous Transaction 123

List of Tables

Table 1.1
Hierarchy of Protocols on the Internet... 6
Table 1.2
Regional Growth on the Internet 7
Table 1.3
Structure of Information Markets.. 8
Table 2.1
Cryptographic Tools & Uses 20
Table 2.2
Information in a Digital Certificate 31
Table 4.1
Information Available to the Parties In a Check Transaction........................ 63
Table 5.1
Information Available in a Credit Card Transaction................................. 70
Table 5.2
Information Available in a First Virtual Transaction 74
Table 5.3
Information Available in a Transaction Using Secure Sockets Layer 77
Table 5.4
Information Available In a SET Transaction 86
Table 5.5
Fields in the NetBill Protocol Definition.. 88
Table 5.6
Information Available In a NetBill Transaction...................................... 91
Table 5.7
Information Available In an Anonymous Credit Card Transaction.................. 96
Table 6.1
Information Available In a Cash Transaction 105
Table 6.2
Information Available In a Digicash Transaction 108
Table 6.3
Information Available to the Parties In a Digicash Transaction 112
Table 6.4
Returns to Scale in Minting Money through Hash Collisions...................... 114
Table 6.5
Information Available In a MicroMint Transaction.................................. 116
Table 6.6
Information Available in an Enhanced MicroMint Transaction 117
Table 7.1
Fields in the Anonymous Certified Delivery Protocol 122
Table 7.2
Information Available with Anonymous Certified Delivery......................... 126

Abstract

In this work I examine the conflict between consumer privacy and data availability in electronic commerce systems designed for the Internet. In particular I focus on the relationship between anonymity and reliability. I do not include systems which require that the consumer has dedicated hardware, such as smart card based systems.

I consider a subset of the policies which affect privacy in Internet commerce systems. Thus I focus not only n privacy laws but also on requirements for data availability. I focus on those system which require information from individual transactions. After this consideration I offer suggestions for possible changes in the regulation of retail transactions.

I select a set of Internet commerce protocols and argue that these are representative. These protocols are: Digicash (Chaum, 1985), traceable Digicash (Chaum, 1985), MicroMint (Rivest, 1996), Secure Socket Layer (Freier, 1996), Secure Transactions Technology (Mastercard, 1996), Anonymous Credit Cards (Low, 1993), NetBill (Goradia, 1994), and First Virtual (First Virtual, 1995a). I consider these protocols based on reliability, security, privacy and regulatory fit. The selection of these systems and analysis techniques are described in Chapter 4. Finally, I introduce a certified delivery layer for the provision of the highest degree of atomicity with anonymous currency. After the discussion of regulatory fit, or how well the system provides consumer privacy and data for regulatory purposes, I consider how changes in the regulations could be made to accommodate the protocols.

In order to provide a broad perspective, I close with a consideration of the regulatory proposals from three viewpoints: law enforcement, data marketers and civil libertarians.

Portions of this work have appeared in (Camp, Sirbu and Tygar, 1995) and (Camp, Harkavy, Tygar and Yee, 1996).

This work contains the opinions of the author, and does not reflect the opinions of the United States Postal Service or the US Government.

1 Introduction & Motivation

Electronic commerce includes sending electronic payments over a public network to obtain electronic goods or promises of the delivery of physical goods. Crucial questions in such purchases are: What can customers, merchants, and banks lose on the Internet? Whom must they trust? And who takes the risks?

Answers to these questions vary across the multitude of proposed protocols for electronic commerce on the Internet. However, an examination of a broad range of these protocols makes clear that in electronic commerce, customers can lose both their money and their privacy.

Can customers protect their privacy and their money? Privacy means that the subject of information controls that information. Anonymity means that information has no subject -- that is, identity is not linked to the information. To protect privacy and money, Internet transactions must be secure, reliable and anonymous. Obviously security is necessary for the protection of both user funds and user privacy. But security alone can protect neither.

Unlike surveillance threats, with anonymous currency illegal acts can be simplified. Risks of anonymous currency include transmitting threats and receiving related ransom anonymously, anonymous blackmail, tax evasion, and trivial money-laundering.

Transactions need to be secure as well as reliable. Reliability requires *atomicity* in the Newtonian sense: transactions must fail completely or succeed completely. The traditional technique for achieving atomicity is rollback, where steps are reversed until the most recent consistent state is reached. For example, if a customer’s attempt to transfer funds from checking to savings fails, funds withdrawn from the customer’s checking account are placed back into the customer’s checking account. If one party is anonymous, however, then rollback becomes difficult; how can money be returned to an anonymous spender? If money cannot be returned in an aborted transaction, then it is destroyed, lost, stolen, or duplicated. Transactions in which money is destroyed, lost, stolen, or duplicated are not reliable.

I do not attempt to address every possible risk inherent in electronic commerce. It is already apparent that the advent of electronic funds transfer can magnify the weaknesses of cash control systems (Fischer, 1988; Mayland, 1993) or entail unnecessarily detailed information gathering that threatens individual's privacy laws (Compaine, 1988; Fenner, 1993; Chaves, 1992; Madsen, 1992). In this section I am concerned with consumer loss of privacy rather than consumer loss of funds -- I focus on the loss of consumer funds and consumer privacy in the analyses of systems.

In the following chapters I explore and resolve the conflict between reliability and privacy in Internet commerce. Because this conflict *can* be resolved, I show that the current regulatory structure encourages the violation of consumer privacy unnecessarily. I identify threats to the funds and privacy rights of participants in electronic commerce transactions under various protocols. I offer a protocol which demonstrates that it is possible to provide customer anonymity while protecting the customer, merchant, and bank from common forms of fraud and loss that could result from network failure. Finally, I offer insights into the nature of distributed electronic commerce based on the two-phase commitment implemented in this anonymous, atomic protocol.

1.1 Money, Its Functions, and Electronic Commerce

Before I describe my dissertation more completely, a consideration of two fundamental questions is in order: Why are reliable transactions important? And, what are the properties of a reliable electronic commerce protocol? To answer these questions, I must first address a more basic issue: What is money? Defined by its three elemental functions, money is a store of value, a standard of value, and a medium of exchange (Rubin and Cooter, 1994). Ensuring that electronic commerce maintains money's functions as store and standard of value is not difficult. In contrast, ensuring that electronic commerce maintains money's function as a medium of exchange is difficult. Money as a medium of exchange requires reliability in transactions, and providing transactional reliability in electronic commerce is not trivial.

Money as a store of value requires durable storage. For money to be a store of value, it must not be easily destroyed or created. If money decays or is destroyed in storage, then it obviously does not succeed in storing value. In contrast, hyperinflation illustrates the failure of money as a store of value when it can be too easily created. Under hyperinflation, entire nations are forced to abandon money and return to barter.

Durable storage is not a critical issue in electronic commerce. Unlike physical money, electronic money is merely bits, and thus can be trivially duplicated. Note that this duplication of money differs from the creation of money only when the duplicates cannot be spent, thus ease of duplication is a double-edged sword. Durable money storage is necessary, but since durable storage is simplified in electronic commerce, it is not a critical research issue in a theoretical study of electronic commerce protocols.

Money as a standard of value requires *interoperability*;[1] that is, to serve as a standard of value, any specific form of money must either be itself widely used (a standard), or readily convertible to another form that is widely used. In the electronic environment, interoperability in terms of wide use means that a protocol can be implemented on many and diverse platforms. This type of interoperability is encouraged by open standards. Low requirements for participation in electronic commerce also encourage interoperability through wide use, by expanding the base of possible customers. Restrictions on participation have the reverse effect. For example, the requirement that electronic commerce customers have a credit card (Mastercard, 1996) prohibits the participation of anyone without a credit history and significant income (e.g., students).

Interoperability in terms of convertibility means different vendors' software can exchange data; in electronic commerce, converting money amounts to exchanging data. Agreements to exchange funds can be handled within the business community, as shown by the evolution of the check clearing system. Interoperability is important for the expansion of Internet commerce. However, interoperability is not a critical research issue in the theoretical study of secure electronic commerce protocols, since even systems that are not secure (First Virtual, 1995a) can provide interoperability.

Money as a medium of exchange requires special transactional properties. As a medium of exchange, money must have transactional durability; that is, money must be conserved in transactions, not created or destroyed. Money transactions must be consistent; the amount received by the seller must be the same amount paid by the buyer, with no change in that amount occurring during the transaction.

---

[1]This does not imply interoperability in the software engineering sense.

The transactional properties that enable money to serve as a medium of exchange amount to transactional reliability. And therein lies the answer to my initial question: why are reliable transactions important? Reliable transactions in electronic commerce are important because they are necessary to the proper functioning of electronic money as a medium of exchange.

There remains, then, the second question: what are the properties of a reliable electronic commerce protocol? The study of distributed databases has defined the characteristics of reliable database transactions as atomicity, consistency, isolation and durability. These are known as the *ACID* properties.

Physical transfers of money illustrate the ACID properties of a reliable transaction. ACID properties are innate in exchanges of physical money. Please note that during this, and all future analyses, I take advantage of gender-specific language to simplify my discussion. The customer is assumed female; the merchant male; and the bank neuter. This allows me to use she, he and it without worrying that the reader may confuse the noun referenced by the pronoun.

Consider a customer's handing a dollar bill directly to a merchant. This transaction maintains:

- Atomicity: The dollar bill will not be lost as it leaves the customer’s hand and is transferred to the merchant. There is always exactly one dollar; it is never duplicated or destroyed. If the dollar is dropped, then the customer can pick it up and return the transaction to its previous state.[2]
- Consistency: After the transaction the merchant knows he has one dollar more; the customer knows she has one dollar less. At no point in the transaction is there ever any confusion over who has the dollar.
- Isolation: That dollar bill will not be confused with a previous dollar bill, so the merchant cannot falsely claim failure to have received payment, and the customer cannot escape her obligation to make payment.
- Durability: After any party receives the dollar bill he or she retains the dollar bill until he or she transfers it in another transaction.

None of these simple physical safeguards necessarily holds in an electronic transaction. When a merchant receives an anonymous payment, it is as if the customer threw a dollar bill across a dark room. Whom should the merchant credit with this payment? Who should receive the goods? In this case, the electronic dollar cannot be identified with a specific purchase or purchaser.

In electronic commerce, payment message must travel over an open network, that is not secure, from the customer to the merchant. Without verifiable acknowledgment in the protocol, the customer will not know that the merchant received the payment message. Under the standard transmission control protocol (TCP), a payment may be duplicated when the communications protocol believes the packet containing the payment message was lost on the network. Moreover, a payment message may be destroyed by network failure. If a payment message is lost, delayed, or destroyed, confusion rather than consistency may result.

---

[2] If this seems impossible, consider writing two checks for the same funds. After the first check is written, the first merchant believes he has been paid, and so he has. B ut if a second merchant receives a check for the same funds, and gets to the bank first, the first merchant’s payment has been invalidated. This may work with electronic as well as physical checks. Similarly with electronic cash, the customer could spend the same dollar twice via copies, and the first merchant to the bank gets the payment.

In this dissertation I argue through repeated analyses that the financial transactions as well as database transactions can also be classified as reliable using these properties. In some systems the financial transaction consists of one distributed database transaction, so in this case the application of these concepts is trivial. In other systems a single financial transaction requires multiple database transactions. In this case the failure of individual messages may require state changes in multiple database transactions for the financial transaction to remain atomic, since the scope of the financial transaction includes multiple database transactions. Throughout this work, I am referring to financial transactions unless otherwise noted.

In sum, transactional reliability is not a trivial matter in electronic commerce. Thus, the provision of reliable transactions is a critical research issue in the analysis of electronic commerce protocols, and one I undertake in this dissertation.

Providing customer anonymity is the second major issue I address in this research. In physical exchanges of money, maintaining customer anonymity is trivial. The merchant present at the transaction may gather some information about the customer through direct observation, but no unique identifying information is recorded and stored as a result of the transaction itself, and no identifying information can be correlated with the purchase. In contrast, electronic commerce is fundamentally the manipulation of computerized records. Purchase information, including customer identity, is easily correlated across electronic transactions. Thus maintaining customer anonymity is not trivial in electronic commerce.

Multiple electronic commerce systems that protect customer identity have been proposed. However, these systems, which provide anonymity, cannot also provide reliable transactions, as they lack atomicity. This dissertation demonstrates that anonymity and atomicity are not mutually exclusive. Customers in anonymous, atomic Internet transactions can keep both their privacy and their money.

1.2 The Internet: The Medium of Electronic Commerce

This dissertation focuses on protocols suitable for commerce on the Internet. Why the Internet? To answer that question, I answer three related questions: What is the Internet? Who's out there? Why Internet commerce?

What is the Internet? And who's out there? The Internet began as the ARPANET, a United States government project for connecting scientific research sites. The tools for internetworking computers were developed by scientists and researchers for use in their own nonhierarchical heterogeneous computing environments. The techniques developed were designed for distributed support, using an iterative process which included seeking and considering comments from the user community. The resulting protocols were open, portable, and enabled the entire research community to share information.

Although the ARPANET connected only a couple of hundred computers at that time, it created the core of compatible inter-networked computers that became the Internet. By 1983, all the networks connected to the ARPANET used the same protocols (TCP/IP) for communication. After the release of Berkeley UNIX 4.2, TCP/IP was included in every UNIX workstation. The UNIX standard created a commercial opportunity for network products (Cerf, 1993). Although the vast majority of these machines were not initially connected to what we now know as the Internet, the ability to inter-network networks became a standard feature for high-end operating systems.

In 1986 ARPANET became NSFNET, and its mission expanded to include students and libraries as well as researchers. In 1990 the first commercial email provider, MCI Mail, was connected to NSFNET. Also in the nineties the National Science Foundation began to reduce subsidies, and gave the responsibility of the NSF backbone to commercial providers, thus enabling a commercial Internet without the limitations borne of Federal funding. Along with commercial email providers, commercial information providers came onto the Internet. Early adopters of Internet technology for information marketing include Dow Jones and Dialog (Cerf, 1993). Thus began Internet commerce.

By 1990 the growth of the Internet was too profitable to be ignored by information providers. However, the market remained primarily technical individuals, with access to information requiring either some understanding of UNIX or proprietary software provided by an Internet service provider (ISP). The growth of the Internet since that time illustrates that the user community has expanded, as shown in Figure 1.1. (The data used for Figure 1.1 came from Internet Domain Survey, 1995a.)

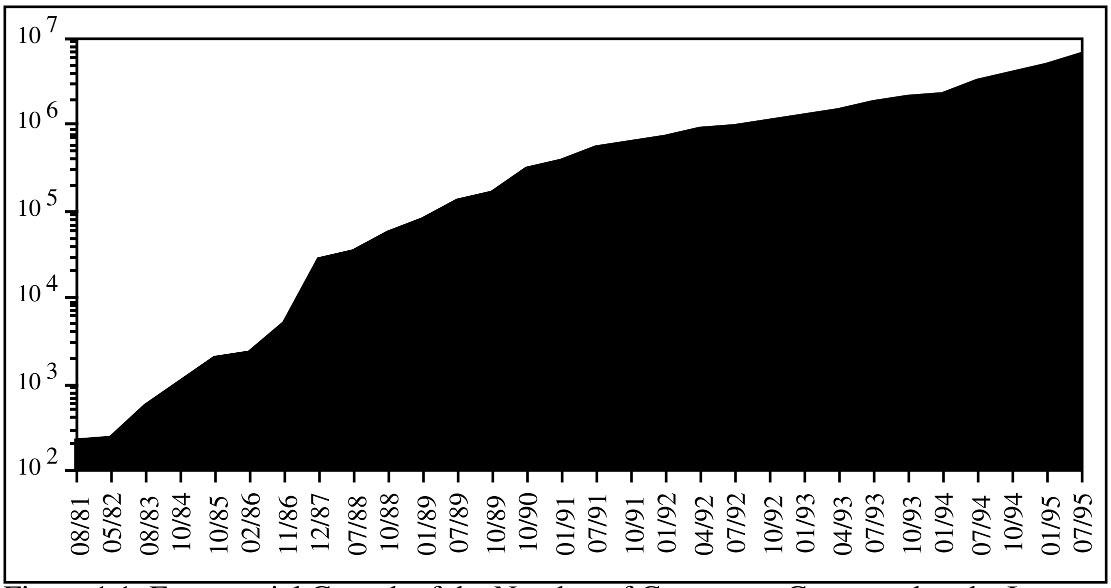

Figure 1.1: Exponential Growth of the Number of Computers Connected to the Internet

A year before the connection of MCI Mail, a European researcher, Tim Berners-Lee, became concerned with effectively transporting the images, postscript files, ASCII text and data files necessary for collaborative physics throughout Europe. The protocol he developed for collaborative physics is the underlying technology for the World Wide Web. The Web allows consumers to search for information on the Internet with a straightforward graphical interface. Easy access to information has been one driver of Web growth.

Protocol	Connects	By Providing
Internet commerce Protocols	Consumer to Merchant	payment, possible delivery verification
Hypertext Transport Protocol	Application to Application	location and presentation
Transmission Control Protocol	Machine to Machine	reliable delivery of multiple packets
Internet Protocol	Network to Network	delivery of packets between networks

Table 1.1: Hierarchy of Protocols on the Internet

With the Web, the Internet became fully capable of supporting user-friendly distributed commerce, just as previous protocols had enabled functionality from simple communication to file transmission. Table 1.1 illustrates how Internet commerce protocols have built on previous protocols, which had in turn expanded the pool of possible merchants and consumers. Of course, Internet commerce does not depend on the hypertext transport protocol (http), as some protocols include options for users with only email.

The World Wide Web is a critical element in emerging markets. Although the Internet began as a specialized US Government project, it is now global. The Internet domain survey has expanded to include ninety countries. The growth of hosts on seven continents from the Internet Domain Survey (Internet Domain Survey, 1995b) is shown in Table 1.2.

The customer base on the Internet grows as the number of countries and connections grows: exponentially with time. Although the coefficient varies across the continents, the form of the curve remains the same. It is these growth curves that so excite the providers of content and commerce services.

Region	Hosts in January 94	Hosts in July 94	Hosts in October 94	Hosts in January 95	Annualized Growth Rate
North America	1,685,715	2,177,396	2,685,929	3,372,551	0.359
Europe, West	550,933	730,429	850,993	1,039,192	0.331
Europe, East	19,867	27,800	32,951	46,125	0.456
Middle East	6,946	8,871	10,383	13,776	0.331
Africa	10,951	15,595	21,041	27,130	0.484
Asia	81,355	111,278	127,569	151,773	0.331
Pacific	113,482	142,353	154,473	192,390	0.268

Table 1.2: Regional Growth on the Internet

1.3 Business Advantages of Internet Commerce

Why Internet commerce? Certainly the obvious answer is, "That's where the customers are." The other answer is that Internet commerce offers the potential to greatly reduce transactional overhead. Many successful business ventures are now on the Internet. Table 1.3 shows examples of businesses on the Internet and corresponding paper information markets (Goradia et. al., 1994).

The Internet supports a range of business functions, not simply payment. Every transaction has multiple phases: discovery, price negotiation, final selection, payment, delivery, and dispute resolution. The Internet can support many types and all stages of Internet commerce (Sirbu and Tygar, 1995).

Product discovery is enabled on the Internet through advertising and electronic word of mouth. Product information is dispersed through Web pages, distribution lists and Usenet groups. The Web enables individuals to locate specific information and search by product or company name. Corporate Web sites often exist solely for the purpose of distributing product information with a simple graphical interface. With distribution lists, or dlists, individuals who have a common interest form a closed group and transmit messages of interest to all members of this group. Announcements of new products are made by members of the distribution list. Usually distribution lists are motivated by discussion, with product announcements being a small fraction of the traffic.

Market Structure	Electronic Example	Paper Example
Publisher pays	WWW catalogs	Mail order catalogs
Advertiser pays	Lycos, Yahoo	Free weekly papers
Club pays	Clarinet, Site license software	Corporate library
Customer subscription	Web magazines, dlist	Professional magazines
Customer pay per item	First Virtual	Storefront sales
Customer pay for time	AOL, CompuServe	Rental items
Mixed ads & customer payment	Prodigy, Netscape business sites	Newspaper

Table 1.3: Structure of Information Markets

Product discovery is enabled on the Internet through advertising and electronic word of mouth. Product information is dispersed through Web pages, distribution lists and Usenet groups. The Web enables individuals to locate specific information and search by product or company name. Corporate Web sites often exist solely for the purpose of distributing product information with a simple graphical interface. With distribution lists, or dlists, individuals who have a common interest form a closed group and transmit messages of interest to all members of this group. Announcements of new products are made by members of the distribution list. Usually distribution lists are motivated by discussion, with product announcements being a small fraction of the traffic.

In Usenet groups new products are announced by subscribers, as is the case with distribution lists. The difference is that Usenet groups are open forums. Not only are product announcements overwhelmed by discussion, but also the information in the groups is notoriously unreliable. Furthermore, direct advertising across Usenet groups is considered offensive by Internet users.

Distribution lists, Usenet groups and the Web overlap. URL[3]'s are sent over distribution list and posted on Usenet, and Web sites connect to archives of Usenet groups and discussion lists.

All the technologies consumers use to find out about services can also be used to locate suppliers. Web search engines, such as the World Wide Web Worm and Lycos, provide a simple way for consumers with Web browsers to locate products.

Price negotiation is supported by email and electronic data interchange. Information goods can be delivered on-line. Customer support can be offered on-line through email and via Web pages.

Every phase of a commercial transaction has associated costs. The ability of an Internet commerce protocol to reduce transaction costs depends on its ability to address these costs. For comparison, the distribution of costs in a credit card transaction is shown in Figure 2.1 below (Sirbu and Tygar, 1995).

The value of Internet commerce partially depends on how the costs in the diagram above can be decreased through automation. The Internet allows administration of customer orders, payment or payment authorization transmission, and production of an invoice to be automated.

---

[3]A URL is a Uniform Resource Locator, i.e. an address for the World Wide Web.

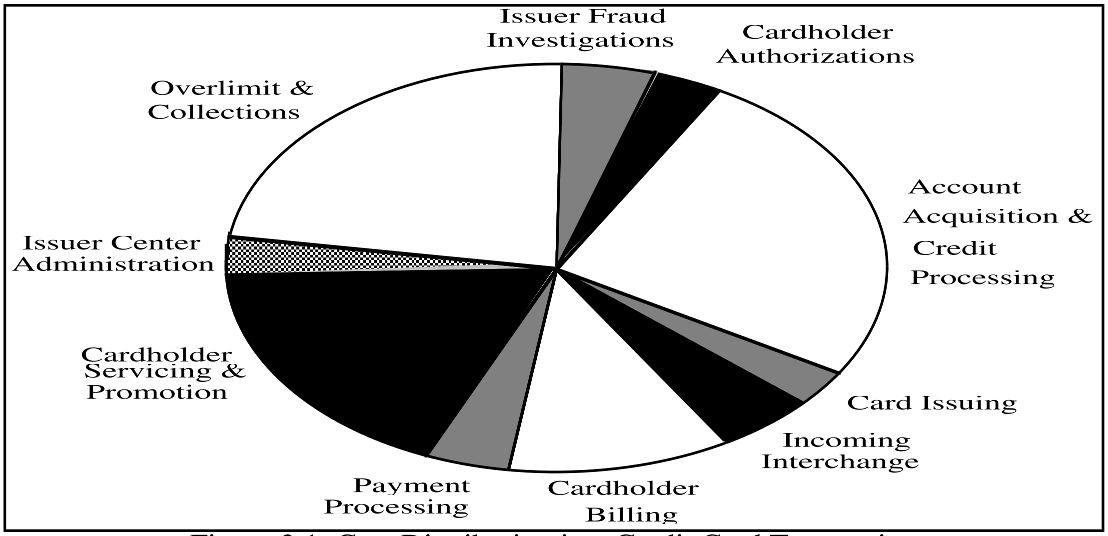

Figure 2.1: Cost Distribution in a Credit Card Transaction

In addition to cost advantages through automation, the Internet allows services to be provided continuously, around the clock, around the globe, in multiple languages, and in multiple currencies. Catalogs of merchandise can be found by interested shoppers at negligible marginal cost to the merchant. The catalogs seen by every consumer can be updated immediately as prices and inventory changes.

Internet commerce could affect the lives of millions. The standards which determine how money and information flow around the Internet are being determined now -- and some of the fundamental decisions about the risks consumers will take are integrated as technical details in technical specifications. Examination of those specifications and enumeration of the risk is particularly timely while Internet commerce is yet infant and the standards still in flux.

1.4 The Organization of this Dissertation

My approach to the investigation of Internet commerce can be most simply described as: define, categorize, select, and solve. In this chapter I have begun the process of definition, with considerations of money and transactional characteristics.

In the second chapter I define the fundamental concepts of security and reliability. These definitions are necessary for the analysis in the following chapters. In the third chapter I continue the process of definition by enumerating the policy constraints on anonymity and information availability in electronic commerce systems. While it is not my position to make a final and exclusive definition of privacy, I do offer multiple definitions from the perspectives of security, jurisprudence and ethics.

In the third chapter I begin the process of categorization. Two concerns of electronic currency are clear: data surveillance and anonymity. The risk of data surveillance is that a ubiquitous commerce infrastructure is built, so that all financial actions are taken in full view of employers, marketers, and government. The Privacy Protection Commission identified electronic commerce as a particular surveillance threat. The report of the Commission stated that an electronic funds transfer system operated by the government would be “an unparalleled threat to personal privacy” and “a highly effective tool for keeping track of people and enforcing ‘correct’ behavior” (Privacy Protection Commission, 1977).

Making systems secure does not solve the conflict between privacy and accountability. For example, security can be used to prevent victims of privacy violations from learning about the information source. Anonymity is a technologically achievable mechanism for providing privacy; yet anonymity is neither security nor privacy. The conflict between privacy and accountability is clear in the conflicting legal requirements for financial transactions: there exist both constraints on and requirements for disclosure.

The entire assortment of statutory and regulatory constraints that can apply to electronic commerce is too broad to be addressed in this work. Although regulatory compliance is often achieved through strictly technical means, laws and regulations typically have one or more social purposes. Underlying motivations range from providing capital for preferred purposes to preventing money laundering. Thus, I classify laws first in terms of their expressed goals and then in terms of the technical means for achieving these goals.

After defining the policy constraints and the technical terms, I select a subset of the currently available electronic commerce systems to examine. In the fourth chapter I describe both my methods for selecting systems, and the methodology for the analysis of the reliability, privacy, security, and regulatory fit of each system.

I first separate the systems according to whether they are *token* or *notation* systems. In token systems, such as paper or gold money, the value is intrinsically part of the item used for exchange. In notational currency, value exists as a matter of record. In notational currency systems, commerce is implemented with the exchange of instructions to alter such records. Checks are notational currency, for example.

After separating systems on the fundamental basis of currency type, I further subdivide the systems. Token systems are separated according to the anonymity provided to the user; notational systems are separated according to business model.

Both notational business-model type and degree of token anonymity are important to the fundamental requirements for trust in an electronic commerce system. In token systems, anonymity determines who is to be trusted with the ability to take untraceable and therefore deniable financial actions. Notational systems are designed according to the role the system is to play in the financial world: intermediary, trusted intermediary, or acquirer. The ability of a protocol to provide atomicity depends on its assignment of roles: if the electronic financial-services provider does not control the final movement of notational currency, that provider may be unable to provide atomicity.

Other characteristics may be incidental to an electronic commerce system, and can be changed in a given implementation. This fact allows me to select certain systems along the spectrum as examples to explore the relationship between atomicity and anonymity.

The token systems that I analyze and their respective level of anonymity are:
• Digicash: token, complete anonymity
• Digicash with embedded user information: token, conditional anonymity
• Micromint: token, no anonymity

The notational systems I will analyze are:
• First Virtual: bank off-line, transactions are not secure
• Secure Sockets Layer: bank off-line, secure transactions
• NetBill: single bank on-line
• Secure Electronic Technology Specifications: multiple acquirers with on-line presence and mutually respected certificates

• Anonymous Credit Cards: customer, merchant and intermediary banks have on-line presence and are separated for reasons of privacy

The fifth and sixth chapters contain the analysis of token and notational systems, respectively. I first explain how a successful transaction would work in the system under consideration. I identity any failure to provide ACID transactions, including a determination of the degree of atomicity provided by each system. Then I consider the worst possible results in the case of a failure of each message, and the worst case for the failure of each security assumption. I evaluate privacy by detailing what information is available to whom. Finally, I consider whether the protocol can meet regulatory constraints, and how regulations could be changed to improve the balance between data surveillance and data availability for each protocol.

The seventh chapter contains the balance of my original contribution to computer science. In this chapter I prove that anonymity and atomicity are not mutually exclusive through the innovation of an anonymous certified delivery layer which can be used with any Digicash-like anonymous token currency.

Finally, in my conclusions I revisit the policy recommendations made in Chapter 3 and reinforced through my analysis of systems. I highlight the importance of the atomicity and anonymity trade-off. I consider three perspectives on anonymous purchases.

2 Security & Reliability

"Each man is responsible for his own acts and omissions only. If he condones what he reprobates, with a weapon at hand equal to his defense, he is responsible for the results."
Warren, 1890

"Strong cryptography can prevent any given message from being read."
National Research Council, 1996

Appreciating the technical analysis in the following chapters and its implications requires an understanding of security, privacy and reliability in electronic commerce. This chapter will provide the necessary definitions of two technical concepts: security and reliability.

Security and reliability are separate issues. A system without security can offer reliable service to consummate users' fraudulent transactions. Yet security and reliability are related. Security requires reliability. Security can provide authorization, authentication and integrity. Security and reliability can provide availability. Reliable transactions have atomicity, consistency, isolation and availability. Properties of secure and reliable systems are defined in this chapter.

2.1 Security

Security is the control of information. Usually security is the control of information by the owner of the information; as systems become increasingly decentralized, security may mean that the user or the creator of the information can control the information.

Note that security is not privacy. Privacy means that the subject of information can control the information. Thus privacy requires security, since security is control over information. However, security is not sufficient for privacy, since the owner and the subject of information may have very different interests in and uses for the data. In fact, security may preclude privacy by assuring that the subjects of information have neither control nor knowledge of the uses of that information.

Security ensures that authorized parties are properly authenticated and their messages are sent through a network unaltered. Thus in a secure system the origin, content and intended recipients of a message can be ensured. Clearly the ability to ascertain the validity of a message is necessary when the information transmitted consist of promises to pay, merchandise to deliver or confirmation of payment.

In this work I will discuss those security strengths and flaws that are inherent in protocol design. This is not meant to imply that design issues eclipse implementation issues, just as in the physical world, a good design does not guarantee a good outcome. Practical approaches to implementation issues can be found in Garfinkle and Spafford, 1986; Denning, 1982; Department of Defense, 1985; National Computer Security Center, 1990.

2.1.1 Threats to Electronic Information Systems

As in the physical world, security is never absolute. There is no case in which it is perfectly impossible to undermine the security of a system. Thus when I examine an electronic commerce system, I consider each security assumption, and the worst case results if that security assumption is unfounded. It is important when estimating the cost of these breaches in electronic commerce systems to recognize that security breaches, once made, can go undetected for some time. In addition, the physical difficulties and dangers that limit the attraction of repeated robberies and break-ins in the physical world do not exist in the electronic realm.

The difficulty of defeating the security mechanisms in a system is referred to as the *work factor*. The work factor for a system is usually measured in processing time and cost for processing power. A system is considered strong, or a message verifiable, if the cryptography used to protect it has a prohibitively high work factor.

Fundamentally there are three ways to obtain electronic information: copy it during transmission, access it during storage or obtain it from an authorized party. Notice that attacking a networked storage facility does not require a physical presence.

Attacks on data transmission include eavesdropping, replay attacks and cryptanalysis. Eavesdropping is the act of surreptitiously monitoring a communication. A common criminal application of eavesdropping is the theft of calling card numbers as they are typed into publicly visible phones. Once electronic information has been stolen it can be easily and anonymously transferred over the network. The benefit to eavesdropping can be reduced or eliminated by encrypting transmissions.

Replay attacks take advantage of the ease of duplication of information. Merchants can attempt to be paid twice by replaying electronic messages that authorize payment. Similarly, individuals can defraud legitimate users of a system by replaying authentication sequences and Shamir to obtain the ability to authorize illegitimate payments. These problems can be solved by using authentication techniques impervious to replay attacks or by adding information in each transaction to make it unique. Authentication techniques that leak no information, and therefore prevent replay attacks using leaked information, are called *zero-knowledge authentication techniques* (Feige, Fiat and Shamir 1987; Tygar and Yee, 1991). Random information can be added to a communication to make it impervious to replay attacks. Such information is referred to as a *nonce.* Transaction identifiers, challenges and timestamps serve this role in many systems.

Encrypted transmissions can be attacked using cryptanalysis. Cryptanalysis refers to the analysis of encrypted transmissions to break an algorithm or obtain a key. Cryptanalysis can be defeated by using secure algorithms with well-chosen keys. It is not possible to protect against cryptanalysis by using a secret algorithm. In fact, using a proprietary algorithm can be very risky, since such algorithms cannot be subject to widespread review.

Cryptanalysis is also used in attacks on the authentication systems protecting stored data. Such attacks are the electronic equivalent of an attack on the bank's vault. Building a secure server is difficult, and the concentration of valuable data in one virtual location can make the value of a successful assault extremely high. Weaknesses in operating system and windowing environments can undermine apparently secure applications. Finally, unlike the case with a physical vault, a successful attack on a secure server may go undetected.

Secure information is commonly protected by passwords. One form of attack on password protected systems requires that encrypted copies of users’ passwords are available. Attackers then encrypt popular passwords, such as dictionary words, and compare these values to those in the files.

The third method for obtaining electronic information, the subversion of security through the confidence of a trusted party, is in no way unique to electronic commerce. The most that security can offer in such a case is the ability to track the individual that improperly released information.

Finally, there are denial-of-service attacks. Here the attacker does not obtain information, but instead prevents anyone else from obtaining information. These attacks limit the availability of a commerce system, denying access to both merchants and customers. The same threats exist in the physical realm: that someone will damage your business premises or threaten your customers. This type of attack is discussed Section 2.1.4

2.1.2 Basic Cryptographic Tools

Cryptography can provide authentication and integrity if properly implemented. Information protected with cryptography can be transmitted confidentially, dated reliably, signed verifiably, and be simultaneously private and verifiable.

The following diagram shows analog equivalents of what cryptography can provide. Cryptography can provide the envelope to prevent snooping for confidentiality. Cryptography can provide the seal on the envelope to assure the message has not been changed; that is, it provides integrity. Cryptography can provide the signature on the bottom of the letter, providing nonrepudiation and authentication of the sender. Cryptography can also provide the lock on the envelope -- assuring authentication of recipient.

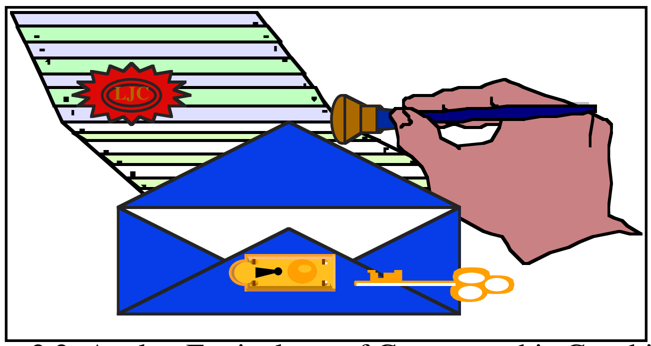

Figure 2.2: Analog Equivalents of Cryptographic Capabilities

Public key cryptography is based on *one-way functions*. A one-way function is something that is easy to do, but hard to undo. For example, it is much easier to multiply two large numbers than it is to factor one large number. The technique in Rivest (Rivest, 1978) is based on the difficulty of factoring numbers. A second common cryptographic authentication technique, Schnorr (Schnorr, 1990), is based on the discrete logarithm problem, while the Feige and Rabin (Feige, Fiat and Shamir, 1987; Rabin, 1978) techniques are based on the difficulty of finding square roots, which is a special case of factoring.

There are two basic types of encryption techniques: public key and private key. In *private key* cryptography the parties wishing to communicate share a single key used both for decrypting and encrypting. If a message is encrypted with this private key, only those possessing the private key can decrypt this message. Because encryption and decryption use the same key, private key cryptography is sometimes called *symmetric*.

Key management in symmetric key systems can be centralized or distributed. It is trivial for users to select a key if there is a central trusted server that will provide a key when requested. However, a centralized key repository is a potential bottleneck and an obvious target for attackers. There also exist key exchange and key generation protocols which allow users to select a key (Diffie and Hellman, 1976). It is easiest for a central server if the users select a key for their own transmissions, as key management can be quite difficult.

With *public key* cryptography each party has a set of keys: a public key and a secret key. Information encrypted with the secret key can be decrypted with the matching public key. This way, secret key encryption can be used as a signature, because the ability to decrypt a document with the published key proves that the owner of the secret key made the original encryption. Information encrypted with the public key can only be decrypted with the secret key. This means that information encrypted with the published key can be widely broadcast but remains unreadable to everyone except the holder of the secret key. Because the possession of one key does not allow you to both encrypt and decrypt messages, public key encryption is sometimes called *asymmetric.*

A third type of useful cryptographic functions is *hash functions*. Hash functions are one way functions with the property that given the output it is difficult to determine the input. With a hash function, information is transformed so that it can be used for verification but not read. The output of a has function is typically much smaller than the input.

Hash functions are subject to an attack based on the *birthday paradox* (Mosteller, 1965). The birthday paradox is this: How many people do you have to have in a room to have a 50% chance that two people have the same birthday? Only 23.[4] This holds across all values -- for example, how many people do you have to have in a room to be 99.99% certain that two people have the same birthday? The answer is 100, although you might guess closer to 365. This is because calculating corresponding birthdays in a room is a special case of sampling with replacement, thus in order to calculate the probability that out of x people, two have the same birthday, the formula is: $1 - \frac{365!}{(365-x)!\ 365^x}$.

This same principle applies when trying to calculate hash value collisions: in a sample of all numbers of x bits, you are trying to find two that have the same value. In order to find hash values, by trail and error alone, an attacker must hash expected $O(2^{n/2})$ values, where n is the size of the hash value. Hash values less than128 bits are not accepted as secure, by many security experts.

For a complete discussion of the algorithms and practical applications of cryptography, see Schneier, 1995.

2.1.3 Availability

System availability can be compromised by malicious hackers, network failures or commercial espionage. Denial of service can be costly, whether resulting from an attack, design failure, or accident. To be useful and marketable, a system must be available.

Availability and scalability can be increased by migrating processing load away from the server to the customer or merchant. This is done with the current network of automated teller machines by requiring the terminal to verify the PIN. In NetBill, the central servers' load is decreased by making the merchant sign using the Rivest Shamir Adleman (RSA) public key system, while the central server uses the Digital Signature Standard (DSS) (Cox, Tygar and Sirbu, 1995).[5]

Availability for the individual merchant or customer is also a function of network availability. If a system depends on real time access, then system availability is a function of the reliability of the network as well as the number and size of messages required by the protocol.

---

[4] This requires two assumptions. First, all births on February 29 are ignored. Second, this assumes a uniform birth rate across all remaining days.

[5] Using both RSA and DSS can serve to distribute load, since DSS signatures require relatively few CPU cycles, but verification is computationally intensive. Conversely, RSA signatures are computationally intensive but easier to verify.

Availability requires reliability, but reliability is not sufficient for availability. Availability means that any system needs to be scalable in the number of users. Availability and scalability are functions of the need for central processing.

2.1.4 Authentication

Secure systems limit the use of resources according to user attributes (usually identity). Authentication establishes user identity or other appropriate user attributes. The appropriate user attribute is then compared against a table of permissions (such as read, write, alter) to determine for which functions the user is authorized.

Authentication is implemented using shared information or the ability to prove unique information. It is most simple to require that one party present the information as proof of identity to another party. PINs and passwords are common examples of simple authentication. This means the presenting party must trust the verifier.

Cryptographic techniques and digital signatures using these techniques enable mutual authentication (Rabin, 1978; Schnorr, 1990; Feige, Fiat and Shamir, 1987; Rivest, Shamir, and Adleman, 1978). With mutual authentication each party can prove authorization to the other and neither party has enough information to later impersonate the other.

In the case of passwords, of which PINs are a special case, the customer's ability to produce a unique number provides authentication. Since the customer gives that number to the merchant's terminal, this means the customer has to trust the terminal. In practical terms, this means when one ATM is badly protected or unreliable, any bank connected to the network can be harmed. This results in attacks such as bogus ATM machines (Davies, 1981; Business Week, 1993; Johnson, 1993), thieves programming cards with others information (Harrison, 1994), and large losses at badly managed machines (New York Times, 1995a; New York Times, 1995b). A similar weakness in the credit card clearing system allows disbarred merchants to use terminals belonging to dishonest merchants (Van Natta, 1995).

The problem of authentication is simplified if there is a mutually trusted authority. This authority can either be on-line to provide verification of authorization upon request, or can provide digital letters of introduction in the form of digitally signed certificates**.**

The problem of untrustworthy hardware can be addressed in three ways: by requiring secure hardware; by requiring the merchants and customers to secure their own terminals; and by accepting the cost of fraud in delivering low-cost items. Electronic transaction systems which require secure hardware are called *off-line* or *smart card* systems. (Smart cards are described in the next section.) Most on-line systems require customers and merchants to assure the reliability of their own hardware. Systems which simply trust the user and accept the corresponding losses are called *crypto-less* systems.

Even when all parties are honest, networks are not always reliable. Therefore, the reliability of acknowledgments should not be critical to an electronic commerce system. With some electronic transactions systems, the protocol assumes reliable acknowledgments. While it is true that high level transactions protocols such as TCP can provide acknowledgments when packets are delivered, there is no acknowledgment of the contents of the packet. Thus the acknowledgments developed for reliable packet transmission are not adequate for verification for electronic commerce transactions. These acknowledgments are not secure; thus they do not provide verifiable information.

2.1.5 Confidentiality

Confidentiality is secrecy. If a message is confidential it can be read only by intended recipients. Eavesdropping is either prohibitively difficult or useless against confidential transmission.

Confidentiality is not security. As in the classic Byzantine Generals problem, if a message saying, "Attack Not, Retreat" became "Attack, Not Retreat" the communications channel has not functioned securely, because of a loss in message integrity.

Confidentiality is not privacy. Gossip is a classic example of confidentially communicated but privacy-violating information. Similarly, electronic information that violates privacy can be transmitted confidentially.

Different degrees of confidentiality are possible. The confidentiality of password protected files depends on the strength of the password selected. Information can be protected with varying degrees of encryption and transmitted over open or closed networks.

2.1.6 Integrity

A recipient of a message with integrity knows that the contents of the message have not been changed. Integrity alone is not security. If a message that claims to be from an account holder is from a thief, integrity does not prevent the theft. Encryption can provide integrity.

A document that is digitally signed is a document that is encrypted. Encrypting a document with a private or symmetric key provides the recipient with some certainty that the document was not altered. Symmetric key encryption provides confidentiality, integrity and possibly authentication. Symmetric key encryption provides confidentiality because only the holders of the symmetric key can read the message. Integrity is provided because when message protected by cryptography is altered it becomes garbage upon decryption. If the key is shared between only two parties, authentication is provided as well, since the recipient knows the sender must have encrypted it.

Using a symmetric key for verification of transmission requires that the recipient and the signer share trust on the contents of the document. With symmetric key encryption, any holder of the symmetric key can modify the document. For this reason digital *signature*s usually refer to public-key signatures, which means that the document is encrypted with the secret key of the sender's public key pair. Public key signatures provide integrity and authentication, and therefore irrefutability. (An action is irrefutable if it can be clearly proven to a third party that the action occurred.) Authentication is provided since only the sender could have encrypted the document with his or her secret key. Integrity results from the cryptographic security of the signature. Since the recipient could prove that the document was encrypted only by the possessor of the private key and that the message has not been altered, public key signatures provide irrefutability. Notice that since anyone with the publicized key can decrypt the message, public key signatures do not provide confidentiality.

Clear signing refers to signing a hash of a document, and sending that with the original document in the clear, i.e. unencrypted. This is particularly effective with large documents because it removes the need for multiple encryption operations. The transmission in the clear of the accompanying message means that clear signing does not provide confidentiality. Clear signing provides integrity, authentication and irrefutability.

2.1.7 Nonrepudiation

Nonrepudiation means that an individual cannot reasonably claim not to have taken an action. Thus, nonrepudiation means an action is irrefutable. In physical commerce nonrepudiation is obtained through controlled hardware tokens (such as credit cards) and physical attributes (like physical signatures).

In electronic commerce nonrepudiation is obtained through use of digital signatures. A digital signature is created when a user encrypts a document using his or her secret key. Then anyone with the user's public key can decrypt the encrypted document and thus prove that the encryption could have been encrypted only by the original user.

In summary, I include Table 2.1 below, which summarizes the uses of various tools and their relationships to the properties described in Section 2.1.4 through Section 2.1.7.

	Zero-knowledge protocols	Hash Functions	Asymmetric Encryption	Symmetric Encryption
Authentication	√		√	√
Confidentiality			√	√
Integrity		√	√	√
Nonrepudiation			√	
Safe from Replay Attacks	√			

Table 2.1: Cryptographic Tools & Uses

2.2 Reliability

Reliable protocols can provide certainty in the face of network failures, memory losses and electronic adversaries. An unreliable electronic commerce system cannot distinguish a communications failure from an attack. If a failure can be used effectively for theft, then certainly such attacks will occur.

Reliability and security are interdependent. The lack of reliability of an electronic commerce system can be exploited by attackers to commit theft. Reliability in electronic commerce may require security to provide authentication, integrity and irrefutability. Reliability is not security. Reliable protocols on servers which are not secure will provide reliable services to attackers as well as authentic users.

Reliable electronic commerce requires fail proof transactions. This fundamental requirement implies other technical requirements. It is widely agreed that an electronic currency system must provide divisibility, scalability in number of users, conservation of money or tamper-resistance, exchangeability or interoperability, and availability (Cross Industry Working Group, 1995; Okamoto and Ohta 1991; Medvinski and Nueman, 1993; Low, Maxemchuck and Paul, 1993; Brands, 1993). The properties which characterize a fail proof transaction are described in this section.

2.2.1 *ACID* Properties

ACID transactions are atomic, consistent, isolated and durable. Distributed ACID transactions are robust and can prevail in the face of network outages, replay attacks, failures of local hardware and errors of human users (Gray and Reuter, 1993).

Transactions are *atomic* in the Newtonian sense; they cannot be split into discrete parts. An atomic transaction either fails completely or succeeds completely. Funds are conserved in an atomic transaction. For example, consider what happens when a customer transfers funds from a savings account to a checking account. Either the checking account is credited and the savings account is debited, or neither account balance changes. There is no case where money either disappears from both accounts or is credited to both accounts.

If a transaction is *consistent*, all relevant parties agree on critical facts of the exchange. If a customer makes a one dollar purchase then the merchant, the customer and the bank (if it is involved) all agree that the customer has one less dollar and the merchant has one more dollar.

Transactions that do not interfere with each other are *isolated*. The result of a set of overlapping transactions must be equivalent to some sequence of those transactions executed in non-concurrent serial order. If a customer makes two one-dollar transactions, then the two payments should not be confused. The customer should not end up being charged twice for one item nor should one single payment be counted twice to give the two dollar total.

When any transaction can recover to its last consistent state, it is *durable*. For example, if the customer physically drops a dollar when making a purchase, that dollar does not disappear. When the customer retrieves the dollar, the last consistent state is restored. Similarly, money that was available to a computer before it crashed should not disappear when the machine reboots.

Atomicity, consistency, durability and isolation in a transaction create the possibility for irrefutability in electronic commerce. Suppose a customer wants to make a purchase from the local software store. The customer must pay, or promise to pay. The merchant either gets payment or proof of intent to pay in a standard purchase order or check. The customer gets a receipt from the merchant indicating that she has paid and expects the merchandise to be delivered. When it is delivered, the customer signs a receipt for the merchant indicating delivery has occurred. Each action is linked with some verification of the action so both parties have some proof in case the other party attempts fraud or fails to perform.

2.2.2 Degrees of Atomicity

Electronic commerce systems have widely varying scopes, some covering only payment while some address everything from negotiation to delivery. Different electronic commerce systems offer different degrees of atomicity to address the problems of remote purchases: money atomicity, goods atomicity and certified delivery (Tygar, 1996).

Of course, electronic transactions may have no atomicity. No atomicity requires mutual trust among participants. The physical equivalent is sending cash or goods in the mail to a post office box. Among electronic currency systems considered here, Digicash has no atomicity. This means that the merchant can claim never to have received payment (Yee, 1994). Customer or merchant fraud can be simple in systems with no atomicity.

Electronic transactions may have money atomicity. The physical equivalent is paying cash. In money-atomic systems there is no mechanism for certification of merchandise delivery. If used for remote purchase with accepted techniques for the delivery of physical goods, money atomicity is quite adequate. But fraud, through a customer's theft of goods or a merchant's refusal to deliver goods after payment, can be trivial when systems with only money atomicity are used for goods with on-line delivery, such as software. Among the systems here, both anonymous credit cards (Low, Maxemchuck and Paul, 1993) and Secure Electronic Transactions have money atomicity (Mastercard, 1996).

Electronic transactions may have goods atomicity. Goods atomicity corresponds to using a certifiable payment mechanism with certified delivery in a physical transaction. Goods atomicity provides high reliability and reduces the opportunity for merchant fraud. Goods atomicity is the equivalent of Collect on Delivery. The merchant is not paid unless there is a delivery. The customer does not pay unless there is a delivery.

Finally, electronic commerce systems may provide certified delivery. With certified delivery the customer only pays if the item delivered matches the description of the item promised. Although this is a semantic clause, it is powerful nonetheless. The merchant is only paid if the item delivered matches the description previously agreed upon by the merchant and customer. NetBill offers certified delivery.

Atomicity depends on design, implementation and business policy. Atomicity depends on funds-available policies because of rollback. Rollback is a technique where all steps are recorded and then reversed until the most recent consistent state is reached. For example, if a customer's attempt to transfer funds from checking to savings fails, funds withdrawn from the customer's checking account are placed back into the customer's checking account.

Rollback is complicated when financial transactions consist of multiple database transactions. For example, suppose a customer orders a frequent flyer award as a free ticket and includes a credit card number to pay for the courier charge. If the entire fare is mistakenly charged to the card, rollback is possible. However, it requires coordinating three databases: the airline frequent flyer database, the airline billing database, and the billing database of the credit card company. This is obviously more complex in computing and organizational overhead than simply re-depositing unused funds at a single institution.

Superficially, electronic transactions are just exchanges of bits, and if the exchange can be reversed, then the transaction can be made atomic. Yet for Internet commerce to expand, there must be some interoperability not only between forms of Internet commerce but also between Internet currency and traditional forms of money. Therefore, if the rollback period is too large the fraudulent party could abscond with unrecoverable cash, making the later acquisition of bits meaningless. This implies that a transaction which implements atomicity using rollback, and that is theoretically atomic, may not be truly atomic. Using two phase commit solves this problem by requiring that the record or funds are locked until global commit is issued. (At the point of global commit, all parties agree that the transaction has been completed.) This implies that for rollback to be enabled for long periods funds should remained locked until commit, so that the money can not be converted in the interim.

2.3 Commerce Transactions

Issues of atomicity and anonymity are complicated by the definition of the scope of a transaction. When does a transaction begin? When does it end? What is the relevant scope of concern in a transaction?

2.3.1 Stages of a Transaction

As I discussed when describing Internet transactions, every transaction has multiple stages, from discovery to dispute resolution. The scope of a transaction limits the capacity of a transaction to provide atomicity. If a protocol considers only the transmission of payment, then discussions of atomicity will arguably be biased against that protocol.

I consider transactions to begin when the customer obtains the means of payment or the customer initiates contact with a merchant. The second is for the case where the means of payment are obtained off-line.

It is my contention that discussion of atomicity is appropriate for every protocol, just as discussion of anonymity is appropriate for every protocol. From the perspective of the customer, if money is stolen there has been theft. If goods are lost, there has been failure. To discuss every protocol only according to the definition of a transaction as provided by its designers would be of limited service: for policy considerations it is appropriate to consider the entire transaction, and not limit the discussion to the framing provided by the designers.

Recall the stages of a transaction:
1. account acquisition
2. browsing or discovery
3. price negotiation
4. payment
5. merchandise delivery
6. dispute resolution
7. collections and final settlement

I will consider all these stages of a transaction for each commerce protocol. Most of the protocols do not include all of these stages explicitly. In many ways comparing Internet protocols is like comparing apples to oranges. Yet such comparisons need to be made for consumers deciding between very different commerce protocols. Thus using consistent language, notation and transactional scope is itself a subtle but real contribution to the debate.

For purposes of commerce it is worthwhile to consider every stage of the transaction. However, I will consider the cost of account acquisition only in the case where account acquisition creates significant expense in a low cost protocol, or where it results in information disclosure in a high-privacy protocol.

In all other cases I will assume the transaction begins with discovery. Both for the sake of consistency, and to reflect the strongest interest in electronic commerce research, I will focus on discovery through the Web. Thus every transaction begins with the information that can be obtained through standard http requests and responses.

2.3.2 Browsing Information

Since I assume a transaction begins with discovery, information which can be exchanged during discovery is properly classified as transactional information. If the customer uses the Internet for discovery, the merchant can obtain information about the customer as she looks through the merchant's wares. This is the first information exchanged in a transaction, and is common to all Internet commerce transactions.

Most purchases are preceded by discovery. A customer cannot purchase an item unless its existence is known. In the analog world, stores obtain information about customer preferences by observing their browsing patterns and set up displays accordingly, or use data on various customers' purchasing habits to target catalogs. Electronic merchants will no doubt do the same and can obtain electronic browsing information easily.

The customer must know of the existence and location of the merchant. There is no need for the customer to know the exact identity of the merchant, merely the host electronic address, unless the commerce protocol requires that the customer trust the merchant.

The amount of information which can be obtained by a merchant during discovery depends on the policies, practices and physical configuration of the customer's Internet service provider (ISP). Other factors which can affect the information available to the merchant include the configuration of the customer's system, the services provided by the Internet service provider, and the software used to access the Internet.

When the customer's client contacts a merchant's server, whether by ftp or Web browser, the merchant can capture the Internet protocol (IP) address of the client. There are intermediaries, called remailers or anonymizers, that can prevent the release of consumer information. However, usually the merchant can capture the customer's IP address.

If the customer's system is protected by her employer's firewall, the IP address may identify only the employer. If the customer is going through the shared IP address of an Internet service provider (ISP), only the identity of the ISP is available. The physical configuration of an ISP may prevent any information but the identity of the ISP from being available -- for example if an ISP dynamically assigns IP addresses as customers access the Internet, as does MCIMail, then only the identity of the ISP is known. For example, for users of American OnLine the only information available is that the customer is using a popular Internet service provider. Thus the minimum information the merchant will have is the name of the customer's Internet access provider, whether that provider be employer, place of learning, or commercial Internet services provider.

However, the customer may use an ISP that provides automatic user identification services. Or the ISP may have a configuration that results in a unique name for each user's personal machine. In these cases the merchant may be able to identify the customer by accessing a daemon, such as fingerid, on the customer's machine. Depending on the naming scheme, various user attributes, such as departmental affiliation (if the machine is in a university setting) will be available (ex. cs.cmu.edu or epp.cmu.edu). In fact, if many users from one company contact the merchant the merchant may build a map of the company's internal network and corresponding users. If the user is from an unprotected site and has a personal machine, the IP address is equivalent to the customer's identity. So the most information passed could be the customer's identity, employer and business area.

Depending on configuration, business area data might not necessarily impart corporate role or job title information with much certainty, since secretaries and senior officers alike have personal machines in the corporate world.

An observer can also detect that there is traffic between the merchant and customer's servers. If browsing information is unencrypted, a well placed observer can watch the merchant's business and have probabilistic information about a customer's browsing habits.

Internet service providers may well provide information aggregated across their user base in marketing information. If this is indeed the case then the merchant gets probabilistic information about a consumer's attributes from the identity of the ISP.

With an IP address, merchant can obtain the customer's hostname using the Domain Name System (DNS), which provides a mapping between domain names (ex., miami.epp.cmu.edu.) and the corresponding network addresses (ex., 128.2.58.26).

If the customer is using the World Wide Web, information availability depends upon the type and version of the customer's browser. With any browser the browser type and version will be sent to the merchant. Netscape also sends the operating system, computer type and helper applications. Depending on the version of the browser used, information including email address and previously visited pages can be obtained by the merchant. The primary purpose of this data exchange is to obtain information about the helper functions and capacities of the client machine to accept various images. This illustrates the interaction between privacy and service.[6]

For a customer to effectively hide identity, the customer must also use an anonymity-providing service to prevent browser-based network information from providing identity information. For an illustration of this, visit http://anonymizer.cs.cmu.edu/. Even with protection from disclosure of browsing information, if the customer is using a single user machine that supports finger daemons then the merchant can still obtain customer identity.

For more effective communications, clients can send information on available helper applications to merchants. Helper applications offer probabilistic information about the consumer’s machine and even interests. For example, the number and variety of helper applications, presence of shareware or freeware applications and the presence of advanced helper applications together imply a level of user technical sophistication.

After sending a request command from the http server, the customer's client will send an accept command. This command can include information on: monitor quality, including size and identification of color monitors, helper applications available, and the quality of the connection. Alternatively, the accept command may just request, “send what you have,” and let the client machine sort the data.

Many electronic commerce protocols begin with an exchange of certificates to assure identity and exchange keys. The customer’s certificate includes the customer’s identifier, the issuer’s identifier and certificate policies. The issuer’s identifier can support the customer’s claim to be authorized on a selected payment method. The qualifiers and certificate policy identifier may provide information on the customer. For example, it may indicate the customer’s credit worthiness or identify that the customer is a student. The dates that the certificate is valid may indicate the customer’s shopping habits or credit-worthiness, based on the policies of the issuer. (Certificates are discussed in more detail in Section 2.5.)

2.4 Secure Hardware

Most electronic commerce systems require secure storage for cryptographic keys, and secure calculations for the consumer, the merchant and some service provider. Even First Virtual, which claims to use no cryptographic assumptions about merchant or client security, needs to have a secure server for its own records.

---

[6]For example, until the release version of Netscape 2.0, servers could obtain the unique email of the user simply by having the user obtain images for their server via ftp instead of http. When using anonymous ftp it has been the polite tradition for the customer’s client to log in as anonymous and then offer her email address as the password. Netscape incorporated this tradition into its browser, so that any anonymous ftp request would include the email address of the user. However, any server could initiate an ftp request without the knowledge of the client simply by preceding the images to be shown in the page with ftp instead of http. The anchors A Href = http://www.cs.cmu.edu/afs/cs/user/jeanc/www/Addie_Walter.gif and A Href = ftp://www.cs.cmu.edu/afs/cs/user/jeanc/www/Addie_Walter.gif would both result in the same lovely image being shown to the customer -- however, the latter would send the customer’s email address to the server.

All systems use secure servers. Some systems also use secure clients in the form of smart cards. In fact secure servers can also be distributed as smart cards. Regardless of the implementation, every electronic commerce system requires secure servers. These servers are obvious points of attacks, as well as potential critical failure nodes. Thus any system analysis should include risk analysis for the case of a failure of a secure server. The design of electronic commerce systems should reflect rules for good security practices for both the security and the financial communities. From the security, each cryptographic key or set of keys should be used for only one purpose so that the subversion of a commerce system should require more than one set of keys. From the financial perspective, subversion of a commerce system should require the cooperation of more than one person, implying that knowledge of cryptographic keys is distributed.

2.4.1 Secure Servers

Security in on-line systems depends to different degrees upon the security of trusted servers. Every commerce system uses servers trusted to some degree: servers that maintain accounts, create electronic money, or authorize transfers. These servers must be both secure and highly available to the customer base.

Although the problem of secure servers is far from trivial, it is simplified by the fact that these are special purpose servers. There is no need for a trusted electronic commerce server to use the most handy tools for potential intruders: mail, ftp and telnet. Furthermore, analysis of security claims is an advanced and maturing field (Davies, 1981; National Computer Security Center, 1985; Denning, 1982).

Although all electronic commerce systems depend to some extent on secure servers, the damage done when a server is subverted varies. In some systems the subverted server could be used to electronically print untraceable money indistinguishable from valid money (Chaum, 1985; Chaum, 1992). There are systems where the subverted server could effectively only electronically print the equivalent of marked bills, because the credits could be detected as false later (Low, Maxemchuk and Paul, 1993; Sirbu and Tygar, 1995). Note that those bills printed after the server was subverted can be identified if the subversion is detected and there are feedback mechanisms that can help identify the time of subversion.

The scope of security needs to be minimized by design. In at least one system, the failure of an authorization server would result only in claims which could be refused by the customer (First Virtual, 1995a). Another approach, as in the secure Web browser provided by Netscape (Freier, Karlton and Kocher, 1996) requires that every merchant's server must be secure. Netscape requires that merchants decrypt and store a credit card number for the time between authorization and settlement to obtain payment, unless the acquirer supports host-based capture. Netscape does not require merchants to delete credit card numbers. Thus any failure of a merchant server could release a large number of credit card numbers.

Having a secure server does not guarantee that the client connecting to the server is secure. The application and operating system must be complemented by the physical security of any trusted device. If a customer or merchant leaves his or her account open, authorized and connected in a public cluster, then the security of the electronic commerce system is damaged despite the best design of server and application.

The hazards of accepting popular software applications as secure is illustrated by three increasingly effective attacks against the Version 2 of the secure Web browsing product from Netscape, Secure Sockets Layer. The first attack against Netscape was simply an illustration of what was already known: any forty bit key is not secure from a determined brute force attack. The second attack was more surprising, and illustrated that the use of predictable information made it possible to obtain Netscape keys in thirty seconds or less (Markoff, 1995). A particularly hazardous attack identified a bug in all Netscape servers that would allow any hostile server to take over any clients (Sandberg, 1995). (This implies that all merchants' servers need to be trustworthy.)

Secure applications for electronic commerce are a matter of both design and implementation. Furthermore, the producer of the consumer application cannot control the computing environment chosen by a customer. Therefore, as in the case of centralized trusted servers, the damage possible in the case of a subverted client application must be balanced by the security of the client application and its environment.

2.4.2 Smart Cards

Smart cards offer the promise of providing every customer with a secure client, at a cost of less than ten dollars and on a scale that can fit in the customer's pocket.

The consumer's machine is almost certain to be the weakest link in any electronic commerce system. This can be addressed with secure co-processors attached to a machine (Yee, 1994) or by giving consumers smart cards. The former requires that each consumer have a dedicated machine; the latter requires that each machine have a card-reader.

A *smart card* is a credit card-sized device for use in a desktop machine or a public terminal. Smart cards are feasible, both in the form of additions to standard computing hardware and as stand alone smart cards. The quality of such hardware varies widely: specialized hardware can be tamper-proof or trivial to defeat.

A smart card provides a trusted computing base where a consumer can store cryptographic keys, and make calculations such as signature without fear of exposing the keys. Smart cards are active devices. This means they can refuse attempts at reprogramming, initiate dialogues, and reject information requests. Because smart cards use secure hardware, providing anonymity is straightforward in debit-based systems, although smart card systems are not always anonymous (Clark and Acey, 1995).

Smart cards can be useful for electronic commerce (Tunstall, 1989; Reid and Madam, 1989; Daper, 1989; Chaum, 1994). Smart cards can resolve conflicts between access to information and privacy by using trusted observers with secure hardware (Brickell, Gemmell and Kravitz, 1995; Brands, 1993). Smart cards can provide information for dispute resolution by providing tamper-proof storage. Smart cards also simplify key management, because they provide secure certificate storage.

Currently major credit card producers are considering adding smart card technology to standard credit cards to increase security (Echikson, 1994; Hansell, 1995; Britt, 1994; O'Keefe, 1994, LaPlante, 1994). Both the addition of secure smart card technology to standard credit card protocols and the failure of Mondex to fulfill its claims of anonymity illustrates that privacy and security are separate issues. With smart cards, spending authorization can be limited to a single transaction, so that traditional attacks based on obtaining credit card or calling card numbers would be fruitless.

Smart cards offer tremendous promise for secure electronic commerce. However, smart cards require the creation of an infrastructure. Millions of consumers must have cards replaced; every merchant must have a card reader.

Smart cards will play an increasingly important role in retail commerce. Yet for commerce over open networks every terminal must have a smart card reader. Alternatively, every terminal could be for a single user with a secure co-processor. Thus any system which requires smart cards essentially requires the construction of a new infrastructure. In my dissertation I have addressed only systems designed for commerce with the current open system that has become this nation's infrastructure.

Note that systems which do not require smart cards are not necessarily incompatible with smart cards. Of the systems considered here only First Virtual is incompatible -- and then because the business assumptions eschew cryptography, not because the protocol is not portable.

2.5 Key Management

Given the current state of cryptography, the problems of security, authentication and confidentiality could all be solved in a straightforward manner if the distribution of cryptography keys were elementary. Unfortunately, this is not the case.

Cryptography can provide confidentiality, authentication, integrity and nonrepudiation only if the keys are protected. This means that keys must be sufficiently large to resist brute force attacks, must represent the parties that they should represent, and must be accessible only by those authorized.

Key management is a critical element of risk management in electronic commerce. The loss of a key should be both unlikely and have limited potential for damage.

2.5.1 Symmetric Key Management

With asymmetric keys the problem of key management is exacerbated by the fact that for every possible set of people, there must a be a unique key. This means for a group of k people to use symmetric cryptography, there must be $\frac{k \bullet (k-1)}{2}$ pairs of keys. These keys need to be changed at regular intervals to minimize the threat of cryptanalysis or the damage possible from a lost key.

Given that one is using symmetric keys, how long should the keys be? Breaking a key is a function of time and money. The price/speed ratio for key breaking is linear (Schneier, 1995). Thus, key length depends on how long a secret must be kept and the value of the information once broken. Key length is not a function of the duration of the transaction since information may be stored long after the transaction is over. If a consumer wants to encrypt a credit card number that will still be useful in ten years, then she would select a key that will withstand a brute force attack funded by an amount equal to the consumer's credit limit for ten years. Factor in the decrease in cost of processing power. For example if the credit limit is $10,000, the card expires in five years and the consumer is using DES (Federal Bureau of Standards, 1977), then the consumer would want a key large enough that spending $100,000 would not break the encryption in five years: i.e. an eighty bit key[7]. Note that the factor of ten is a result of the decrease in the price of computing power resulting from Moore's Law[8] (Schneier, 1995). Since DES is a block cipher, keys must be 56 bits or a multiple of 65 bits, so an eighty bit minimum implies a 112 bit key.

---

[7]Cryptography using 40 bit keys can be exported under current Federal regulations. Greater key lengths may be exported only with key escrow under current proposals. Thus, in practice a consumer gets 40 bit protection.

[8] Moore's Law notes that the density of silicon integrated circuits has closely followed the curve

Of course, in practice there are off-line controls. Thus instead of a credit limit the consumer might want to substitute \$50, since the consumer could only lose \$50 if the credit card number was stolen. However, some party will take the loss for unauthorized credit card use, and the maximum value of that loss provides a conservative estimate for key length. The fact that the transaction took only one minute is not a guideline for key generation.

Given that key length has been determined, how does one distribute the keys? By definition there is not already a shared key. Simply sending the key would provide no security, since any observer of future transactions would have a copy of the symmetric key.

This problem has been addressed using a common trusted entity who can generate keys and already has independently authenticated both parties. In practice this would be the bank or financial services provided for a commerce protocol. However, having a central party run the key exchange is a serious bottleneck, even given the existence of a trusted third party. The most commonly used algorithm for obtaining keys from a trusted third party is Kerberos.

Symmetric keys can be obtained if the users in question already have asymmetric keys.

2.5.2 Asymmetric Key Management

The first asymmetric algorithm invented, Diffie-Hellman, was created for the exchange of symmetric keys by users who have no shared trusted party. In fact, given any asymmetric key algorithm, key exchange is trivial -- the initiator of the conversation encrypts the desired key in the receiver's public key and signs it with her own key. This would provide confidentiality and integrity. It would further provide authentication if the symmetric key was linked to the initiator.

Thus with asymmetric keys, distribution of the key itself is trivial, while linking a key to an individual is difficult. It is simple for me to post a public key and claim it is mine. Presumably after that anything sent in that key could be readable only by me. Of course, it is equally simple for me to post a key and claim it belongs to William Jefferson Clinton. Thus the issue in asymmetric key management is linking the public key to the individual.

Key length is also an issue in public key systems. When attempting to break asymmetric keys the attacker attempts to factor the number that is the public part of the key set. However, it is reasonable to compare asymmetric key lengths for systems based on factoring and symmetric key lengths based on the difficulty of brute force attacks. Considering only brute force attacks, 56 and 112 bit DES keys are roughly equivalent in strength to 384 bit and 1794 bit RSA keys, respectively (Schneier, 1995).

There are two basic philosophies in the distribution of asymmetric keys: hierarchical and non-hierarchical. The non-hierarchical system is used by a system called Pretty Good Privacy (Zimmerman, 1995). The hierarchical system is used by most designers of electronic commerce systems. Hierarchical key management systems for general use have been proposed by the United States Postal Service (The Economist, 1996), and Verisign (Verisign, 1996).

(bits per square inch) = $2^{(t - 1962)}$ where t is time in years; that is, the amount of information storable on a given amount of silicon has roughly doubled every year since the technology was invented.

With Pretty Good Privacy, a user publishes his or her key, and other users can endorse this key. First a user generates a key. Then the user endorses that key, so that some other person cannot claim the key[9]. The user publicizes the public key. Other users then endorse the key by signing a hash of the key.

Public key endorsements create a "Web of Trust" that takes advantage of off-line relationships and reputation. There is no single hierarchy that can verify every user for every situation -- only a set of people vouching for one's goodwill. This creates a network where a person offers her reputation for proof that a key links to an individual. Thus, once you establish a reputation on a Usenet group your endorsement will be meaningful on that group. However, if your reputation has been established on misc.kids the endorsement is meaningless on alt.cyberpunk. Interestingly enough there has been no monetary market for key endorsements.

The meaning of a key endorsement is only that the endorser believes that the holder of the key is, in fact, who he claims to be. There is no implication that the endorsed party is supported by the endorsers as being trustworthy on any other count. There is no implication that the endorsed party likes the endorsing party. There is no implication of honor, agreement or trustworthiness.

Thus in a Web of Trust each person has limited power to state that an individual is linked to a key. Each additional signature is a probabilistic increase in proof, with some signatures more trustworthy than others.

Alternatively, a single source can be determined to have complete power in stating that a key corresponds to an individual. This trusted party may verify others as having the power to connect individuals to keys; however every key/identity link is based on the trust of the first party. This is a hierarchical system, and the single trusted party is the *root.*

The mutual trusted party can provide digitally signed electronic credentials suitable for off-line authentication. These credentials verify that ownership of a public key pair corresponds to an attribute, usually identity. Multiple parties are planning to operate public key hierarchies. Competitors in the market for the provision of electronic credentials for electronic commerce include Verisign, Banker's Trust and the United States Postal Service (Verisign, 1995; The Economist, 1996).

Certificates can be used to connect an individual to any attribute, such as a person to a public key. Examples of off-line certificates include credit cards, drivers license and club membership cards. Just as one person holds many off-line certificates, one person can hold multiple on-line certificates.

In practice a certificate in public commerce both links an individual, an attribute and a public key. For example, a Secure Electronic Transactions certificate links a consumer with an identity, the right to authorize a charge against a Visa account, and the public key used to verify a payment authorization. The credential may contain a pseudonymous account number (PAN) instead of an account number. Visa and Mastercard consider the certificate in SET to be the electronic representation of the bank card.

With a certificate, key management concerns are: the length of the root key, the length of individual keys, the number of roots, and the lifetime of the certificate. For a certificate to be valid it has to have integrity and authentication must be possible. A signature from the

[9]Such an attacker could not read your mail or sign documents in your name. However, such an attacker could implement a denial of service attack.

root or any trusted authority can provide both of these given that the root key is secure and the information in the certificate is still valid at time of use.

Thus the loss of security for the secret root key of a certificate chain results in all certificates becoming suspect. If the secret root key is compromised, the attacker can create or alter certificates; for example, inserting a bogus public key set into an otherwise valid certificate and thereby obtaining the ability to authorize payments on another's account.

X.509 is the dominant standard for certificates. The required fields in X.509 are shown in Table 2.2.

The distribution of certificates allows the trusted third party to provide off-line verification of multiple attributes. The distribution of the certificate also implies distribution of identity and association information.

Field	Purpose
Version	version 1,2 or 3
SerialNumber	unique (within Issuer) serial number, assigned by Issuer
Signature	algorithm used to sign the certificate
Issuer	trusted entity which signed the certificate
Validity	dates between which the certificate is valid
Subject	identity of the valid holder of the certificate
SubjectPublicKeyInfo .algorithm .algorithmIdentifier	algorithms for which this certificate is valid
SubjectPublicKeyInfo .subjectPublicKey	public key of the holder of the certificate
IssuerUniqueID	unique identifier of trusted entity
SubjectUniqueID	unique identifier of holder of the certificate, assigned by trusted entity
Extensions.extnId	identifies extensions (version 3 only)
Extensions.critical	Boolean, use described in .extnId above
Extensions.extnValue	extension data

Table 2.2: Information in a Digital Certificate

An otherwise valid certificate can be used by an attacker if the associated secret key has been compromised. It is fairly easy to obtain a certificate since certificates are public affirmations of attributes. It should be quite difficult to obtain a secret key. Thus for a certificate to be trustworthy, the secret key corresponding to the certificate must be secure. Thus in the analysis of systems in Chapters 4, 5, and 6, a secret key compromise is considered the crucial security breach.

A certificate may be used fraudulently if the information or the attributes attested to in the certificate are incorrect. This may result because of fraud when the certificate was issued, or if information changes after the certificate was issued (such as the loss of credit privileges) (Simpson, 1996).

When certificates are renewed keys should be changed. Thus in the case of private key compromise the ability to commit fraud ends with the lifetime of the certificate. This suggests that key lifetime has a significant effect on fraud. However, if certificate lifetime is too short the cost of certificate issuance may outweigh the benefits of fraud reduction. See Simpson, 1996 for a discussion of optimal certificate lifetime.

2.6 Microdata Security

Microdata security is the protection of the identity or attributes of individuals. Microdata security can be compromised by obtaining information about a specific entity either from data sets where such specific queries are prohibited, or from the correlation of information across data sets.

In electronic commerce systems personal data is distributed transaction by transaction. Information on any one transaction usually provides very limited information about a consumer. However, compiled transactional data can provide a detailed summary of a person's habits, which offers information about the consumer's habits, preferences, income and possibly beliefs. Such small data elements which can be compiled to create a threat to privacy are called *microdata*.

Microdata security focuses on disclosure. A disclosure is not necessarily a violation and a violation of privacy may not be a disclosure. For example, TRW, a credit information agency, collects data about purchasing patterns of individuals both to provide credit references for consumers upon request and to market that information. The first purpose is a service to the consumer when the consumer is trying to obtain credit. The second may be a privacy violation. Both uses of the data are disclosures of consumer information.

Microdata security is concerned with all types of disclosure without consideration of intent. There are four types of microdata disclosure: identity disclosure, attribute disclosure, inferential disclosure and population disclosure (Duncan and Lambert, 1989).

Identity disclosure is the release of information clearly associated with an individual. A University's release of the Social Security numbers of its students would constitute identity disclosure.

When linking a record with an individual provides additional information about that individual, then attribute disclosure has occurred. As an example, the Social Security numbers of the students mentioned above may allow anyone receiving that information to obtain the students' credit, employment, or medical history. Attribute disclosure has been a primary concern in the widespread use of Social Security Numbers or any other universal identifier.

Inferential disclosure is the release of information which does not identify associated individuals. This does not mean data have no unique identifiers, only that those identifiers cannot be linked to specific people. For example, the records of the New Haven needle exchange program require unique identifiers, but there is no way to identify the person who is exchanging needles as only code names are used (Kaplan, 1991; Kaylin, 1992). The concern over inferential disclosure is that given a set of attributes, identity disclosure may occur.

Population disclosure is the release of information associated with a defined population. Population disclosure can result in privacy violations depending on the size of the population and the information released. The release of the mean and standard deviation of the salaries of five employees would be an example of population disclosure resulting in a high risk of privacy violation due to sample size. A well known example of population disclosure of innately sensitive material was the release of the name of a high school with a high HIV positive rate. This left students at the high school subject to harassment (McGraw, 1992). The third concern with population disclosure is that repeated population disclosure can result in attribute or identity disclosure. An example of repeated disclosures that can violate privacy is the release of the average salary in an institution immediately before and after one person joins the staff.

The risk of disclosure from the release of data is difficult to calculate. This risk can be reduced but not eliminated through masking (Duncan and Lambert, 1986). Methods of masking include: adding bogus records which leave aggregate values unchanged; switching values among different respondents; and removing identifiers. However, a significant risk of population and inferential disclosure may remain after masking.

Studies of disclosure offer useful definitions and methods for developing privacy-protecting systems. Microdata security offers insight into the threats to privacy that can result from a data compilation. The microdata security paradigm recognizes the different threats to privacy created by compilations of different types of data and identifies some vulnerabilities of data compilations to privacy-violating misuse. The study of microdata security clearly illustrates that the release of a single element of information must be considered in the context of all other possible data releases and not as an isolated incident. The application to transactional records is clear.

2.7 Psuedonymity & Anonymity

The analysis of microdata offers the possibility of obtaining information about one's travels, beliefs, financial status and any current medical conditions. However, this information can only be collected if it is possible to link information from a transaction to an individual. Two ways to prevent such linkage are pseudonyms and anonymity. (Much of this overview is based on Froomkin, 1995.)

Despite its critical role in privacy, identity is merely another data field in an electronic information system. Any information may be hidden during a transaction. When the information that is hidden is the identity of the customer, then that transaction is *anonymous*. Identity includes any user attribute that can be easily linked to exactly one individual: user id and domain name, Social Security Number or IP address of a single-user machine.

Anonymity means that the identity of a party cannot be determined during or after a transaction. Conditional anonymity means that a party cannot be determined during a transaction, but may be determined afterwards with the cooperation of one or more record-keeping parties. True anonymity is technically feasible for electronic commerce, but for reasons of law enforcement such anonymity may not be desirable. In fact, anonymity is illegal for some transactions in many jurisdictions, including the United States.

Pseudonyms are aliases. Pseudonyms may provide continuity in an otherwise anonymous environment, or they may be a special case of conditional anonymity. Psuedonymity means that a customer can be identified by a pseudonym during a specific transactions or set of transactions, but the user’s identity cannot be determined. A pseudonym may provide authorization or identify certain attributes (for a discount for repeated use, for example). A user may choose to have unique pseudonyms for each transaction, to use the same pseudonym for multiple transactions, or to have a pseudonym for each merchant. Without a delivery address, or with an intermediary that hides the delivery address, a pseudonym provides no identity information.

Traceable psuedonymity means that the chosen alias can be linked to the user's true identity. Many so-called anonymous remailers really provide traceable pseudonyms, since the records of the remailer can reveal the identity of the user of the service.

Credentials usually link an identity to an attribute. Credentials can also be used pseudonymously. Unlike a pseudonym alone, credentials by definition provide membership information, thereby giving partial identity information. For example, a user of a CMU discount is one of seven thousand, not just one of millions.

Credentials can in fact be used anonymously to link two attributes, ex. the right to authorize a purchase and asymmetric keys for verifying and creating a purchase order. However, this use is proposed first in this dissertation in Chapter 7.

The value of a pseudonym in terms of privacy protection is a function of its frequency, duration and breadth of its use. A pseudonym used many times in multiple situations becomes equivalent to identity. For example, the use of Social Security Numbers has become so common that a Social Security Number is now equivalent to identity; and like identity, Social Security Numbers are linked to many attributes.

In the absence of legal protection, anonymity offers consumers the only protection against data surveillance. Unfortunately, wide spread availability and use of anonymity has its own dangers. A recipient of an anonymous electronic threat knows this truth too well.

A detailed mechanism for anonymous and pseudonymous communication was proposed by Cox, 1994.

3 Internet Commerce, Privacy and the Law

"The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized."
Fourth Amendment to the Constitution of the United States

"The Fourth Amendment demands that we temper our efforts to apprehend criminals with a concern for the impact on our fundamental liberties of the methods we use."
Florida v Riley, 1989, J. Brennan, dissenting

There is a fundamental conflict in every information system, including electronic commerce systems: privacy versus data availability. This conflict also exists in the law. Currently there are legal requirements both protecting and prohibiting privacy. In this chapter I identify specific laws and general principles that affect privacy and are likely to affect electronic commerce. These include laws which impose limitations or requirements on consumer information processing.

There is a broad philosophical base underlying the concept of privacy. On that base are built state laws, Federal laws, and constitutional prohibitions. The applications of these laws to information technology, and digital commerce in particular, are as yet undetermined. Information technology will require a new balance between information privacy and information disclosure. Information technology enhances the availability of data, making information easier to collect, analyze and disclose. Conversely, information technology also enables people to better protect their privacy. Cryptographic techniques in particular can protect privacy.

At times, the laws and regulations on privacy appear to be a puzzle, constructed of particular laws for each category of information. Under which category or categories the Internet will fall is far from clear. Thus, here I simply outline the structure of privacy law at a high level, and then consider examples from the most relevant categories. To do so, I begin with a brief overview of state laws. The state laws are based on tort law, which is the earliest privacy protection in the American tradition, drawing heavily on common law. After a discussion of state laws, I consider Federal statutory law, and finally, constitutional law.

Having discussed privacy, I turn to data availability. I consider Federal law on information technology and financial transactions, as well as Federal cryptography policy. Reporting requirements of financial transactions and limitations on strong cryptography are the regulations that most affect the design of Internet commerce systems. Each of these topics is considered in a separate section.

In conclusion I offer proposals which address the need for a new balance between information availability and privacy in the American regulatory regime.

Please note that this chapter focuses on the law in the United States. This is because the United States is likely to be a leader in regulating Internet technologies, partially for the historical reasons briefly described in Section 1.2. Other nations have different legal traditions, principles and social norms which may cause these nations to choose a different point of equilibrium between privacy and data availability. Reaching international consensus of the relative value of privacy and information availability will undoubtedly be problematic.

3.1 Privacy

Although Americans have long valued privacy, the law has been somewhat slow to recognize a right of privacy as such. In a now famous law review article, Warren and Brandeis made an eloquent case for recognition of a legal right to privacy (Warren and Brandeis, 1890). Warren and Brandeis justified this new legal right by pointing to a number of judicial decisions rendered in different fields of law, and found in these decisions the core idea that privacy is an interest that needed explicit legal protection. Warren and Brandeis advocated the right “to be let alone.”

In lawsuits brought thereafter, litigants relied on the Warren and Brandeis article as authority for the idea of recognizing a separate and new legal right of privacy, interference with which could be remedied even in the absence of a statute or constitutional provision explicitly protecting privacy interests. A number of state courts found such arguments persuasive. Through common law (that is, case-by-case) developments in state courts, the privacy rights of Americans were recognized and protected in a wide variety of situations.

Although much of the legal protection of privacy interests remains a matter for state common law, some states have also passed privacy protection statutes. Some statutes, such as New York's, are of a general character; others, such as those that protect the confidentiality of library records, are very specific. A few states also have constitutional provisions that protect privacy.

The Federal government has generally been less involved in privacy law than state courts and legislatures in large part due to a general Congressional inclination -- one that has constitutional overtones because of the limited powers the U.S. Constitution confers on Congress -- to leave to state law the legal protection of personal interests. Nevertheless, in furtherance of its powers to promote interstate commerce and communications, Congress has enacted a number of special laws, such as the Electronic Communications Privacy Act and the Right to Financial Privacy Act, that protect privacy.

Furthermore, through a series of cases interpreting the bill of rights provisions of the U.S. Constitution, Federal courts have come to recognize in the First, Third, Fourth, Fifth, Ninth and Fourteenth Amendments to the Constitution the basis for inferring a more general constitutional right to privacy. While the best-known of these decisions is Roe v Wade, which announced a constitutional right of privacy in relation to decisions about whether to seek an abortion, there are a number of constitutional privacy decisions.

Despite this long history, privacy remains a difficult concept to define with precision. This difficulty is compounded by the fact that the law of privacy develops state by state, in response to particular disputes brought to the courts. Furthermore, some privacy interests continue to be protected under the rubric of separate legal doctrines, most notably that of trade secrecy law. Finally, privacy rights are of two distinct types: rights of autonomy and rights of solitude.

Privacy rights are also not absolute. They often conflict and must be reconciled with other social, economic or legal interests, such as the right to speak freely, even about others. Even the United Nations Universal Declaration of Human Rights (United Nations, 1995) defines a limited privacy, recognizing only the right to be free from unwarranted intrusions, rather than all intrusions[10]. In addition, many industry groups have lobbied against more extensive privacy legislation, arguing that privacy interests are better protected through more flexible and consensual efforts such as industry adoption of codes of fair information practices.

Thus, to provide a complete overview of privacy I begin with a consideration of Codes of Ethics which have been offered in the absence of law. I then discuss state law and Federal law. At the Federal level, I consider statutory and constitutional law separately.

[10]Article 12 of the United Nations Universal Declaration of Human Rights states, “No one shall be subject to arbitrary interference with his privacy.” By comparison, the same document recognizes an individual's right to life with no qualifiers.

3.1.1 Codes of Ethics

Various organizations have attempted to address privacy outside of the legal structure, through codes of ethics. Some believe codes of ethics to be the single most powerful deterrent to abuse of computerized data (Hoffman and Clark, 1991). I examine ethical codes covering privacy in access to electronic information, from three sources: libraries, computer scientists, and the Code of Fair Information Practice.

The code of ethics of the ACM illustrates computer scientists' position on privacy. General Imperative 1.8 of the ACM Code is explicit: Respect the privacy of others. The accompanying guideline states: “This imperative implies that only the necessary amount of personal information be collected in a system, that retention and disposal periods for that information be clearly defined and enforced, and that personal information gathered for a specific purpose not be used for any other purpose without consent of the individuals.” (Anderson, Johnson, Gottenbarn and Perrolle, 1993). Other imperatives which address privacy require that computer professionals honor confidentiality (Imperative 1.9); system managers evaluate the potential risks of any computer system (Imperative 2.5); and all ACM members articulate and support policies that protect the dignity of every person affected by a computing system (Imperative 3.5).

In practice, computer scientists must grapple with the technological difficulties imposed by privacy. In the case of electronic commerce systems, designers sometimes view privacy as incompatible with reliability. Providing a reliable system that includes privacy protection embedded as anonymity, is problematic at best. Thus fulfillment of the imperatives for privacy is difficult in highly reliable electronic commerce systems, and is often limited in system design.

Libraries have a tradition of advocacy for civil liberties. That tradition is reflected in a strong code of ethics, one libraries have respected in practice. In 1938 the American Library Association extended its Code of Ethics to obligate librarians to “treat as confidential any private information.” Libraries are historically among the largest information distributors and processors in a community, although theirs has been an analog tradition.

All systems of records could conceivably be guided by the Code of Fair Information Practice. The Code of Fair Information Practice provides a baseline of privacy-protecting practices which can be used to compare electronic commerce systems. The Code states that data compilations should never be secret. Furthermore, for existing data compilations, there must be a way for individuals included in the compilation to find out what information is stored about them and how the information is used. The Code also requires that individuals have the ability to audit and correct their information. A mechanism for the individual to prevent disclosure is required by the Code; however, the prevention of disclosure is clearly identified as the responsibility of the organization which controls access to the data.

The Code of Fair Information Practice prohibits secret data compilations; that is, the code requires that subjects included in data compilations should know that the compilations exist. This requirement illustrates how security and privacy are orthogonal: a subject would be unable to access his or her own information in a very secure system. Subjects of the compilation may not even know of its existence or their inclusion, since by definition the existence and contents of secure information are not widely disclosed.

3.1.2 State Law

The mere existence of a right to privacy that is universally recognized through the UN, in this nation's legal structure and in common law, is important. If privacy is a right, those who gather and disseminate data, not the individual, bear the responsibility for maintaining privacy. However, this right has not been implemented in information technology: privacy is part of ethical codes but not consistently part of computer code. Privacy protection through codes of ethics has proven inadequate (Office of Technology Assessment, 1985; Office of Technology Assessment, 1986; National Research Council, 1996).

State law is primarily tort law. Tort law is civil law. In criminal offenses the state is the prosecuting agent. By definition, criminal acts are offenses against the state. In civil cases, two parties argue the case and the state serves as the impartial agent for judgment. Civil law is distinguished from criminal law in that it concerns only those violations that can be addressed with monetary compensation; only the state has the right to ask for imprisonment.

Trying to make conceptual sense of a plethora of common law privacy cases, Prosser in his 1941 treatise on tort law identified four kinds of privacy rights cases: intrusion upon seclusion, appropriation of name and likeness, false light, and public disclosure of private facts. While some have challenged the appropriateness of this categorization (Halpern 1991; Bloustein, 1968; Kalven, 1966), the separation of privacy violations into four separate torts is the judicial standard. The cases I cite come primarily from Alderman and Kennedy, 1995; Trublow, 1991; and Speiser, Krause and Gans, 1991.

Intrusion upon seclusion is the violation of the right to be let alone. The original definition of privacy clearly singled out the press for intruding into private affairs: “Gossip is no longer the resource of the idle and of the vicious, but has become a trade which is pursued with industry as well as effrontery.” (Warren and Brandeis, 1890). But what is electronic seclusion? Is it one’s own electronic mailbox where particular messages are unwelcome?

Appropriation of name and likeness is the use of a person’s name, reputation or image without his or her consent. An early and well-known case is that of a young woman who found her image distributed throughout the city on bags of flour. She had given no consent and received no compensation. The makers of the flour had thought her face would be commercially useful and that she was owed no compensation for the luck of having such a countenance. The New York courts agreed. Despite her failure in seeking restitution, appropriation of name and likeness is universally recognized when there is commercial gain. There are limits in different states on the ability to seek restitution if there is no commercial gain.

False light is the publication of information which is misleading, and thus shows an individual in a false light. This is similar to libel. The ability to charge another with false light depends on the standing of the victim and the role of the privacy violator. Private persons (as opposed to public figures) need show only falsehood in a case of false light; however, concerns over speech rights hinder the pursuit of restitution. False light is recognized as an offense in all fifty states, under one rubric or another. However, some states do not recognize misrepresentation as a privacy violation per se.

Public disclosure of private facts is self-explanatory. Information deemed as “newsworthy” can be printed even if it is a violation of privacy. Currently public disclosure of private facts is treated seriously in some jurisdictions, including New York and South Carolina. However, in some jurisdictions, notably North Carolina and Texas, one cannot bring action under this tort (Alderman & Kennedy, 1995).

One recurring question concerns how much privacy people forfeit when they are the victims of crime or require emergency assistance. A new twist on this question is: how much privacy does one forfeit when one becomes electronically active? Do Internet posts make one validly subject to other posts which disclosure private facts? With this, as with other torts, the results will vary between states.

In addition to tort and case law, states offer statutory and constitutional protection of privacy. The level of protection varies widely among states.

Ten states[11] have privacy as an explicit right in their constitutions. Of those, only Louisiana and California provide relief to public sector employees, while the other amendments deal exclusively with the ability of the state to obtain information. Application of state constitutional law on electronic information remains undetermined.

State laws vary with respect to the categories of electronic information which are protected. In the next paragraphs, I briefly enumerate the states that offer protection and the classes of information for which the protection is offered.

States with specific protection of financial transaction information are: Alaska, Arkansas, California, Florida, Illinois, Kentucky, Massachusetts, Minnesota, Mississippi, Missouri, Montana, Nebraska, Pennsylvania, Texas, and Wisconsin (Trublow, 1991). The laws in Arkansas, Massachusetts and Montana apply only to records of electronic funds transfer. The protections in other states limit disclosure of consumer financial information.

In addition, fourteen states protect all financial information, not merely transaction-specific data, from state governments: Alabama, California, Connecticut, Illinois, Louisiana, Maine, Maryland, Massachusetts, Montana, Nebraska, Nevada, New Hampshire, Oklahoma and Oregon (Trublow, 1991). These laws are similar to the Right to Financial Privacy Act at the Federal level (which is discussed later in this chapter.)

Forty-one[12] states and the District of Columbia also have specific statutes on the confidentiality of library circulation records. The District of Columbia also provides statutory protection for library records. The importance of this protection is that some of the records of purchases for information goods on the Internet will provide information on a consumer's regular reading habits, much like library records.

State also may protect information in electronic form. State statutes that may protect privacy include both wiretapping statutes and broad computer crime statutes. Computer crime statutes at the state level often focus on manipulation of financial information for fraudulent purposes, and thus resemble the Federal Wire Fraud Act more than the Federal Computer Fraud and Abuse Act. (Both of these acts are discussed in detail later in this chapter.) States which offer specific protection against abuse of computerized financial information for personal gain are Arizona, California, Delaware, Georgia, Michigan, New Mexico, North Carolina, Tennessee, Texas, Utah and Virginia (Nimmer, 1992).

---

[11] Alaska, Arizona, California, Florida, Hawaii, Illinois, Louisiana, Montana, South Carolina, and Washington.

[12] These states are Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Illinois, Indiana, Iowa, Kansas, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Missouri, Montana, Nebraska, Nevada, New Jersey, New Mexico, New York, North Carolina, North Dakota, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, Vermont, Virginia, Washington, Wisconsin, and Wyoming (Johnson, B.S., 1989).

Under wiretap laws states may protect telephone numbers, which provide electronic location information analogous to IP address on the Internet. For example, in California and Pennsylvania, courts have ruled that telephone numbers as offered by Caller ID have some protection under wiretapping statutes. Yet the reach of wiretapping statutes into other electronic realms can be limited. Again consider California, where the courts have ruled that intercepting email is not wiretapping, and have requested that the legislature clarify the issue (Trublow, 1992).

A particularly problematic issue in the application of state law is the nature of location in electronic information systems. Suppose a customer has a credit card account in Wisconsin, an ISP in her home state of California, and makes a purchase from an electronic merchant in Delaware. If the purchase information is intercepted and compiled for internal use by a company in Utah, where did the interception take place? Was it a wire tap? Is it judged under the local jurisdiction of the credit card headquarters, as would be the case if the customer was concerned with usury? Does it matter if the company in Utah is taking a demographic survey of the customers of the Delaware merchant? If the company in Utah makes no money but is trying to make marketing decisions about general Internet purchasing habits, is that wire fraud or legitimate use? None of these questions is simple; and they are further complicated by uncertainty of jurisdiction.

3.1.3 Federal Law

At the Federal level there exist special statutes to protect privacy and constitutional guarantees of privacy. These are separately addressed in this section.

3.1.3.1 Statutory Law

Four areas of Federal law can apply to information created in electronic commerce: controls on financial information; controls on electronic information; controls on the investigation of citizen access to information; and laws enabling the regulation of cryptography. The confluence of consumer privacy and cryptography is a recent technologically-driven event, and is addressed in a separate section. Laws concerning financial and electronic information and citizen access to information are addressed here.

The Privacy Act of 1974 codifies the principles of the Code of Fair Information Practice (as discussed in Section 3.1.1) and requires its practices be followed for all government databases and databases of government contractors. The Privacy Act requires that individuals be informed of all government compilations of data of which they may be part and limits the sharing of data between Federal agencies. The Privacy Act also limits the use of Social Security Numbers as universal identifiers in Federal databases.

Specific protections exist for various classes of personal information in analog forms including medical, video rental, criminal and financial records. The electronic purchase of information falls not only under the umbrella of constitutionally protected free and private access to information (see Section 3.1.3.2); it also falls in the category of financial exchange. Laws covering privacy of access to information have a different tradition than laws governing commercial information; these laws are based on the assumption that financial records belong to the bank and not the consumer.

The Right to Financial Privacy Act was enacted in response to a Supreme Court decision, one that denied rather than defined privacy: United States v Miller (United States v Miller, 1976). United States v Miller determined that there are no Fourth Amendment constraints limiting government access to one's financial records. There is no expectation of privacy in bank records under the Constitution.

The Right to Financial Privacy Act extends the Fourth Amendment and the Code of Fair Information Practice to bank records. It limits the conditions under which any institution can disclose customer information to Federal authorities. Yet the Right to Financial Privacy is not as complete as Fourth Amendment protection because of broad exceptions. The Right to Financial Privacy Act includes exceptions when the bank is acting in its own self-interest, for regulatory proceedings, IRS summonses, and required reports. The Right to Financial Privacy Act also does not apply to aggregate information. The Right to Privacy Act offers limited protection; for example, the court has ruled that the admission of financial records which have been stolen from a third party by the contrivances of a government agency are admissible in court (United States v Payner, 1980).

The Fair Credit Reporting Act applies the principles of the Code of Fair Information Practice to credit reporting agencies. Unfortunately, this act applies only to those agencies which provide credit reports as their *primary business function*. This means that financial information given to credit card companies, banks and other institutions can be freely traded without consumer knowledge or consent. However, the Fair Credit Reporting Act has been effective in preventing the proliferation of private credit guides which contain information on individuals. Previously, private credit guides offered detailed, if unreliable, information on easily identifiable individuals. Now the only guides allowed are those which contain encoded information, whereby identifying a consumer without the information on a payment instrument is not possible.

The Fair Credit Reporting Act protects the credit agency from the charge of negligent release in the case of misrepresentation by the requester. Credit agencies must ask the requester the purpose of a requested information release, but need make no effort to verify the truth of the requester's assertions. In fact, the courts have ruled that, "The Act clearly does not provide a remedy for an illicit or abusive use of information about consumers" (Henry v Forbes, 1976).

The Fair Debt Collection Practices Act similarly limits dissemination of information about a consumer's financial transactions. It prevents creditors or their agents from disclosing the fact that an individual is in debt to a third party, although it allows creditors and their agents to attempt to obtain information about a debtor's location.

The information in Internet commerce will be both financial and electronic. So, laws which provide protection to electronic information as well as those that provide protection to financial information are worthy of consideration.

Transmission of electronic information is addressed by the Electronic Communications Privacy Act (ECPA). The ECPA establishes criminal sanctions for interception of electronic communication. The statute calls for imprisonment of not more than five years, a fine, or both. Given that these are strong penalties, there are also broad exceptions.

The exceptions to the ECPA are for those who act under the color of law; when the party intercepting the communication is also a party to the communication; or when one party has given prior consent. The first exception refers to law enforcement personnel, and gives access to communications with a warrant. The second exception, when the party intercepting the communication is also a party to the communication, is obvious. A financial agent, such as a bank or credit card company, that is party to a communication has the right to read and reference that communication. In the third exception, prior consent can be explicit or implied. Consent may be implied by an employee agreeing to work in an environment where system use is required or it may be explicit in a written request for financial services. Thus the Electronic Communications Privacy Act provides limited protection from law enforcement, but it does provide legal protection against observers.

The Computer Fraud and Abuse Act of 1986 was an extension of the Counterfeit Access Device and Computer Fraud and Abuse Act of 1984. It extends that act to prohibit six types of conduct. The three prohibitions which are of most interest are: intentionally accessing a computer without authorization and obtaining information in the financial record of a financial institution; knowingly and with intent to defraud, accessing a Federal interest computer and causing damage of more than $1,000; and knowingly and with intent to defraud trafficking in computer passwords.

The definition of Federal interest computers covers more than is readily apparent. For example, a computer of Southwestern Bell at a local exchange office was determined to be a Federal interest computer due to the ubiquitous and critical nature of the telephone system (United States District Court, 1992). As electronic commerce becomes more widespread, the Internet may be classified as a system of Federal interest.

Theft of financial records, theft of services, and removal of data from public sector computers are all explicitly prohibited. Information trespass is made criminal by this act. Furthermore, the Computer Fraud and Abuse Act specifically identifies viewing financial information to be a breach of the law.

Note that only an *unauthorized* violation of privacy is a matter of concern for the Computer Fraud and Abuse Act. If the owner of information, such a mortgage company or medical information clearinghouse, sells the information, there is no abuse or fraud. If information is obtained for one reason and then sold to another party to be used in a fundamentally different way, then there is no fraud or abuse. This is further illustration that security and privacy are disjointed.

The Wire Fraud Act has a target similar to the Computer Fraud and Abuse Act. It is of interest because of its focus on financial information. It prohibits devising a scheme to commit fraud and transmitting signals in order to commit fraud. The Wire Fraud Act has been used to prosecute hackers who access computerized phone systems and reprogram them to obtain free long distance services. Presumably the Wire Fraud Act would be used to prosecute those who use Internet commerce for fraud, but not those that use Internet commerce for surveillance.

Like the Computer Fraud and Abuse Act and the Electronic Privacy Communications Act, the Wire Fraud Act focuses on unauthorized viewing or theft of information goods: the use of consumer information is not an issue.

3.1.3.2 Constitutional Law

States are constrained by the United States Constitution, as well as their own. In 1969 the Supreme Court made the right to privacy explicit in Griswold v Connecticut. The Court found the right to privacy implied in the Constitution in the First, Third, Fourth, Fifth, Ninth and Fourteenth Amendments (Compaine, 1988). The Constitutional right to privacy continues to be recognized by the courts in accordance with the tradition of the particular class of information.

The First Amendment states:

> "Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the government for a redress of grievances."

The privacy implications are that people under surveillance are not likely to express views, go to assemblies or religious meetings with which the agencies of surveillance are likely to disagree. The freedom to read is actually the freedom to read without fear of surveillance. The Court has ruled that the right to privacy covers the right to read --- unobserved --- material which the Federal Government finds objectionable. Specifically, in Lamont v Postmaster General the Court stated that "any addressee is likely to feel some inhibition in sending for literature which Federal officials have condemned."[13] The Court has found a right to privacy in association and political activities. The Court has ruled that the right to privacy covers memberships and personal associations (NAACP v Alabama, 1958), confirming the "right of members to pursue their lawful private interests privately and to associate freely with others."

The Third Amendment states:

> "No soldier shall, in time of peace be quartered in any house, without the consent of the owner, nor in time of war, but in a manner to be prescribed by law."

The Fourth Amendment states:

> "The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized."

Together the Third and Fourth Amendments create a region of privacy, a space inviolate by the government except in constrained circumstances. These Amendments suggest that what one does in one's own home is not the business of the government. Note that members of the NAACP were found to have not only the First Amendment right to associate, but also the right to "pursue private interest privately," as one might in one's own home.

The Fifth Amendment states:

> "No person shall be held to answer for a capital, or otherwise infamous crime, unless on a presentment or indictment of a grand jury, except in cases arising in the land or naval forces, or in the militia, when in actual service in time of war or public danger; nor shall any person be subject for the same offense to be twice put in jeopardy of life or limb; nor shall be compelled in any criminal case to be a witness against himself, nor be deprived of life, liberty, or property, without due process of law; nor shall private property be taken for public use, without just compensation."

The government cannot imprison people without charge, nor require that they speak. The implication is that the government has no right to hear all that you might say, thereby intruding into your thoughts. Just as the government has no right to search your papers by the Fourth Amendment, the government has no right to search your thoughts by the Fifth. Nor does the government have the right to arbitrarily limit your movements.

The Ninth Amendment states:

> "The enumeration in the Constitution, of certain rights, shall not be construed to deny or disparage others retained by the people."

---

[13] The United States Postal Service was required by §305 76 USC 840 (the Postal Service and Federal Employees Salary Act) to detain mail considered "communist political propaganda" and release it only upon the request of the recipient.

Without the Ninth Amendment, the right to privacy could not be found in the Constitution. The right to privacy is nowhere specifically identified in the Constitution. Thus, without the Ninth Amendment's specific identification of the list of rights mentioned as not being exclusive, the right to privacy as implied by the other Amendments could not exist.

Section 1 of the Fourteenth Amendment states:

> "All persons born or naturalized in the United States, and subject to the jurisdiction thereof, are citizens of the United States and of the state wherein they reside. No state shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any state deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws."

Sections 2-5 of the Fourteenth Amendment address apportionment of representatives and Civil War disqualification and debt, and thus are not of interest here.

None of the rights set forth in the Constitution can be abridged by the States. If the Federal Government has no right to your home, speech or papers, neither do the state governments. The rights, which together provide privacy from the Federal Government, provide privacy from state and local governments as well.

The Constitutional right to privacy differs from state tort-based laws in that it is focused on individual autonomy rather than the communications of others. Privacy allows individuals to take certain actions without fear of retribution, rather than prevent the publication of information. In fact, privacy rights prohibiting intrusion into seclusion and publication of private information have been limited at the Federal level precisely because of First Amendment protection of speech rights.

The Court has decided that there is no constitutional right to privacy or expectation of privacy in financial matters (United States v Miller, 1976). Consumer financial information is voluntarily supplied to financial institutions; the information is owned by that institution; and there is no reasonable expectation of privacy for financial information because by its nature it must be shared in the course of business.

Constitutional protections of privacy have not been consistent. Often a delay extends between the introduction of new technologies and the extension of privacy rights to the users of that technology. Consider the case of telephony. In 1928 the Supreme Court determined that no person has a right to privacy in telephone conversations (Olmstead v United States, 1928). The Supreme Court ruled that recording telephone conversations was not a search under the Fourth Amendment because the conversation left the defendant's home on lines which could not be secured. The Court stated that since the technology was inherently without security, people knowingly sacrificed privacy when they communicated using the telephone. The Supreme Court reasoned that telephone correspondents knew that the signals went outside their homes and only the most naive would expect privacy. Olmstead reads: "There was no searching. There was no seizure. The evidence was secured by the use of the sense of hearing and that only. There was no entry of the houses or offices of the defendants. . . . The language of the amendment cannot be extended and expanded to include telephone wires, reaching to the whole world from the defendant's home or office. The intervening wires are not part of his house or office, any more than are the highways along which they are stretched."

The reasoning in Olmstead applies to the Internet today. Of course, this reasoning remains true for the telephone network as well. For the decades between Olmstead v United States and Katz v United States (Katz v United States, 1967), the law of access to telephone conversations essentially stated that because the system was open, privacy was not to be expected. Although there were great technological advances in telephony between 1928 and 1967, the change in the law arguably reflected changes in social rather than technical practices. The Court has not determined which judgment applies to the Internet today.

3.1.4 Privacy & Information Technology

The first definition of privacy in Warren and Brandeis, 1890 was written in response to technological threats to privacy. Specifically, Warren and Brandeis were concerned with the scandalous reporting of the press and its lack of regard for privacy. These gentlemen felt that the new technology upset the previous balance between privacy and the availability of information, thus forcing a reconsideration of the right to privacy. A century later information technology is again changing the balance between privacy and data availability.

Electronic information technology changes the balance between privacy and information availability because electronic information is easy to compile, correlate and distribute. Monitoring every keystroke of users of information technology is trivial. Information, once collected, is easy to analyze and distribute (Turn and Ware, 1976; Pool, 1983; Office of Technology Assessment, 1985, Compaine, 1988; Computer Science and Telecommunications Board, 1994).

Consider the effect of information technology on the four torts. In a practical sense, invasions of privacy were once the purview of the press and government by virtue of the difficulty of publication and surveillance, respectively. Information technology has made eavesdropping and publication easy for all Internet users, thus increasing the opportunity for privacy violations. Currently the lack of authentication or integrity makes the dissemination of false information much easier.

In the case of intrusion upon seclusion, electronic trespass has been defined as a crime at the Federal level. Yet electronic intrusion upon seclusion has yet to be defined. One possible electronic case of intrusion upon seclusion can be found in the beta release of Microsoft's networking software. Early hackers with beta versions of Microsoft network software noted Microsoft would have sent consumers' machine capacities and entire directory structures to Microsoft upon product initiation. After this fact was publicized, Microsoft reduced the amount of information to be transmitted and offered the consumer a choice to "register" with the company. If in the electronic world one's own hard disk is electronic seclusion, then this could have been a tort violation.

Consider the capacity to show in a false light. Certainly the capacity to distribute other people's private communications, through re-transmitting email or building Web pages for general distribution, creates new possibilities for false light. Email may be edited and displayed as "evidence" that an individual has certain beliefs very unlike his or her own. A combination of loaded labels, unidentified sources and hidden agendas can be used to present an issue in a false light (FAIR, 1996). Loaded link names, anonymous email, and misleading domain names make these tools of deception available to everyone on the Internet.

To determine the effect of information technology of the appropriation of name and likeness, one must ask, what is one's electronic likeness? What value must exist for it to be misappropriation? The difficulty of applying appropriation of name and likeness in the electronic realm is illustrated by the search engine, Alta Vista, http://www.altavista.com/. With Alta Vista, DEC has provided the ability to search Usenet postings by author or keyword. DEC is using the speech and ideas, the electronic visage of an Internet user, for the purpose of advertising DEC's Internet acumen. Yet this is an offense against no person in particular and no money changes hands --- the only possible "profit" is an increase in the number of hits to DEC's Alta Vista site.[14]

Finally, consider the fourth tort, public disclosure of private facts. Public disclosure becomes very easy when everyone is a publisher. For example, here at Carnegie Mellon University one student's homosexuality was revealed on an electronic departmental bulletin board. Communicating such a fact to many of the department faculty and all of the student's colleagues was vastly simplified by the use of information technology.

In short, information technology has altered the balance between privacy and data availability by giving many people the power to compile and disclose information, powers that were previously held only by governments and the press.

3.2 Cryptography Policy

Where information technology has increased threats to and breaches of privacy by increasing data availability, it has also increased the power of individuals to maintain their privacy --- particularly through cryptography. However, the Federal prohibition of the export of cryptographic technology (discussed in the following section) has effectively prevented the widespread implementation of strong cryptography in operating systems and communications packages intended for the global market (Froomkin, 1996). This prohibition has affected the design of electronic commerce systems intended for export, including Secure Electronic Transactions and the Secure Sockets Layer.

Cryptographic technology has historically been the purview of governments. In the United States the Coordinating Committee for Multilateral Export Controls explicitly classified cryptographic technologies as exclusively military technologies shortly after the Coordinating Committee's creation in 1949. Control of exports of cryptographic technology falls under the Export Administration Act, the Arms Control Act, and the International Traffic in Arms Regulations.

The view of cryptography as war technology is expressed in the Export Administration Act, which prohibits "the export of goods and technology which would make a significant contribution to the military potential of any other country." Thus cryptography has been controlled as Auxiliary Military Equipment under the International Traffic in Arms Regulations (ITAR).

Under ITAR, cryptographic technology can be exported if restricted for the following purposes: copy protection, authentication, financial information, integrity without confidentiality, compression, or prevention of theft of information services (such as pay television). Thus the export of any device which provides strong cryptography for protecting privacy in an electronic system is prohibited without a specific license.

Note that the recommend key lengths for encryption from Chapter 2 would be allowed for the encryption of payment information in systems intended for international use. However, a general use system, such as the Secure Sockets Layer, cannot provide strong encryption (for export) because it could also be used to provide general communications privacy.

---

[14] A visit to a particular file on a Web server is called a hit. The number of hits on a Web Site is a count of its popularity and a suggestion of its producer's Web-publishing abilities.

Current national cryptographic standards are not adequate for market use (National Research Council, 1996; Schneier, 1995). In 1988, the National Institute of Standards and Technology (NIST) determined that the Federal Information Processing Standards (FIPS) on encryption needed to be replaced, and that any replacement should be “public, unclassified, implementable in both hardware and software, usable by Federal agencies and U.S. based multinationals.” The criteria reflect the needs of Internet commerce: portability across operating systems, exportability across national borders, the need to provide privacy as well as financial security, and the capacity to optimize for specialized applications.

The actual NIST standard fell short of these goals, as have the subsequent proposals. The Clipper chip embodied the first proposal for key escrow as defined in the Escrowed Encryption Standard (National Institute of Standards and Technology, 1994). Due to strong objections[15] the requirement that the Federal Government maintain the databases for key escrow has since been removed. This FIPS has been followed by three more proposals for escrowed systems.

The analysis and evaluation of cryptography policy suggested by the criteria has recently undergone a complete review by the National Research Council. The resulting report, Cryptography’s Role in Securing the Information Society (National Research Council, 1996), recommends a new approach to cryptography. This report’s recommendations most relevant to electronic commerce are:

• National cryptography policy should be developed by the executive and legislative branches on the basis of open public discussion and should be governed by the rule of law.
• National cryptography policy affecting the development of commercial cryptography should be more closely aligned with market forces.
• Export controls on cryptography should be progressively relaxed but not eliminated.
• The U.S. Government should take steps to assist law enforcement and national security to adjust to new technical realities of the information age.

Cryptography policy has managed to appear to be constantly in flux for several years, without many significant changes occurring. Currently there is consideration of a bill in the Senate to lift controls on export, and there is a case under the First Amendment. The Clinton Administration has proposed a new key escrow proposal, popularly called Clipper III. In 1992 it was widely assumed that any White House containing now Vice President Al Gore would be an advocate for freedom to export, export controls remain in place today. Thus while it may appear that this problem will be addressed shortly, there is historical basis for believing that this will not be the case.

Cryptography policy is a case of conflict, with data availability on one side and privacy and security on the other. Cryptography is restricted for the purposes of law enforcement and national security. Removing constraints on cryptography would serve commercial, security, and privacy interests. Controls on cryptography illustrate that there are reasons other than reliability for providing identity information.

---

[15]In the NIST comments period on the Clipper proposal, NIST received 320 comments. Of 22 government, 22 industry and 276 individual comments, 2 were positive, 4 were neutral, and the remainder covered the range of negative, from those expressing specific misgivings tocondemnations of the complete proposal.

3.3 Data Reporting & Disclosure

There exists a conflict between protecting consumer privacy and assuring that information is available to government so that it can perform its legitimate duties. Thus, in addition to laws protecting privacy there are laws requiring data reporting and disclosure.

New technologies for electronic payment present new risks and require new regulatory approaches, while the basic social policy remains unchanged. To consider regulatory requirements for electronic commerce, I strip away existing conventions derived from paper models to categorize the underlying social goals and to suggest new regulatory approaches based on the new mix of concerns raised by electronic commerce.

The fundamental choices for reporting economic data remain when the data becomes electronic. However, financial information is more prone to privacy violations than other information, not only because financial information is innately commercially valuable, but also because the Fourth Amendment does not apply to financial information.

I close this chapter with general policy suggestions. I refer to these suggestions as appropriate in the system analyses of Chapters 5 and 6.

3.3.1 Required Information Reporting

Reporting requirements are a tangential and possibly minuscule part of funds transfer and electronic banking laws. (For the purpose of this chapter, reporting requirements are requirements that information about a transaction or a set of transactions be available to either party of a transaction or to government.) Yet the laws requiring information availability for government can have tremendous impact on the design of electronic commerce systems --- not least by unnecessarily prohibiting the provision of consumer privacy in electronic commerce. Current reporting requirements are biased by the assumption of a paper currency model, where transaction information is either documented on paper or potentially unavailable.

Types of laws requiring information for the private sector and the public sector are both of interest here, especially considering the decreasing distance between public sector and private sector data repositories. For example, in order to track both tax debtors to the Federal government and parents delinquent in providing children support, the Federal government has weakened the Privacy Act and the Social Security Act, by expanding the use of Social Security Numbers as universal identifiers. The laws which require the use of Social Security Numbers as financial universal identifiers are the Debt Collection Act and Child Support Enforcement Act, respectively.

Sometimes the biases against privacy in reporting requirements are purposeful, as with laws which prevent fully anonymous systems. Sometimes, however, there is simply a mismatch between the paper model on which these laws were based and the strengths and weaknesses of electronic currency systems. I hypothesize that within each category of electronic currency system there are technical enhancements that can alleviate the need for the trade-offs that have been necessary in paper currency systems.

In focusing on consumers' financial information, there are two classes of business which should be specifically addressed: depository institutions and consumer credit reporting agencies. These two business types maintain most consumer financial records, and reporting restrictions are often specific to these institutions. Businesses which currently handle the most cash --- credit unions, savings and loans, banks and thrifts --- have specific record-keeping and reporting requirements. These businesses will be discussed together as depository institutions. (Note that other regulatory changes, especially the loosening of controls on line-of-business and fewer marketing restrictions, are allowing these businesses to merge and converge.)

To consider reporting requirements without the biases created by the assumption of paper currency, I separate reporting requirements along two dimensions: system requirements and social goals. First, I consider the range of system requirements that are inherent in regulatory requirements; for example, data must be trapped in a transaction; data must be stored and searchable by account number. Second, I consider the set of underlying social goals that motivate financial reporting requirements. A single unifying principle in regulations on commerce is lacking. Instead there is a set of underlying reasons which together motivate nearly all the reporting requirements.

After having categorized reporting requirements according to underlying social goals and data availability requirements, I construct a matrix which spans the range of goals and techniques. The general alternatives for reporting in the electronic realm are considered. Finally I make specific suggestions for making the techniques more compatible with the capacities of electronic commerce so that either the goals can be better fulfilled, or violations of privacy can be reduced without preventing attainment of the desired goal, or both.

3.3.1.1 Techniques for Regulatory Information Requirements

The specific requirements for information found in the law are manifestations of social goals. Within electronic information systems, there exist a wide range of techniques for assuring information availability, including anonymous updates to aggregates and distributed escrow. With regulations based on a paper model, this same range does not exist -- there are four basic techniques used for obtaining the data necessary for the public sector to fulfill its legitimate purposes. These are:

- immediate reporting
- periodic reporting
- periodic aggregate reporting
- data storage requirements for later access

To implement immediate reporting, a commerce system can filter transactions on some variable, such as purchase amount or item purchased, and then initiate some action when the conditions for immediate reporting are met. Alternatively, the commerce system can prohibit transactions of a given type to avoid immediate reporting requirements, or assume that such reporting is the responsibility of the merchant.

Periodic reporting requires collection of data to be reported for verification of reports. However, the data may be disposed of after the report or deleted from general access computers so that internal misuse is minimized.

Periodic aggregate reporting in the paper world requires that all records be kept for verification of reported values. This practice is followed in the electronic realm as well. However, there are greater options for data reporting in the electronic realm. One such option is encrypting individual records with an escrowed public key and updating the periodic values before the records are securely stored.

Data storage options are also greater in the electronic realm than the paper arena. Data may be stored encrypted or in such a way that subversion of a single depository provides no useful data.

Note that where a requirement fits within these categories is sometimes a matter of interpretation. For example, filing taxes is aggregate information in that all individuals' sources of income are aggregated over the year. It is periodic reporting in that each individual can be tracked by taxpayer identification or Social Security Number. With this caveat in mind consider an example in each of the four reporting categories.

An immediate reporting requirement was initiated by the 1988 Money Laundering Act, which empowered the Treasury to require that all suspicious transactions be recorded. (The Money Laundering Act extended the provisions of the Bank Secrecy Act) The Treasury interpreted this to require that banks report all cash transactions above $10,000 and all purchases of financial instruments (such as traveler's checks) over $3,000.[16] All $10,000 transactions must be reported by all merchants, using the appropriate forms, to the Treasury. This requirement was expanded by the Money Laundering Act of 1994 to include all money transmitters, such as Western Union, American Express, and currency exchange houses. Given the extension in 1994, it appears likely that requirements of the Money Laundering Act will apply to transaction processors in electronic commerce systems.

The most common periodic data reporting requirement is the annual individual Federal tax filing on April 15. Wages, tips, and other forms of income must be reported to the Federal, state and local governments as necessary for tax purposes. The details of expenditures can be reported according to taxpayer preference. That the increased record keeping possible in notational currency systems would be effective in preventing tax fraud is suggested by the 800,000 "dependents" who disappeared from the tax rolls as soon as their Social Security Numbers were required (Davis, 1995).

The Community Reinvestment Act requires periodic aggregate data reporting. The Act requires financial institutions to make credit and depository services available to all the neighborhoods in their service area on an equitable basis. Typically this means loan application aggregates sorted by ethnicity of the borrower, neighborhood, or loan amount. Specific data requirements vary over time, states, and even between institutions.

The Bank Secrecy Act, despite its name, actually limits consumer privacy by requiring detailed record keeping. It was passed to assure law enforcement access to detailed records of personal financial transactions under subpoena. The Bank Secrecy Act requires that financial institutions maintain records of all transactions over $100 for at least five years. Note the Bank Secrecy Act requires not only that records be kept in plain text, but also requires that the record be an image of the bank's record of the transactions, such as a copy of the check. While the Bank Secrecy Act applies only to cash and to cash-like instruments and not to wire transfers at this time, it is reasonable to include the Bank Secrecy Act in this analysis, given the number of systems which use cash and checks as the basis for their model, as illustrated in later chapters.

Many of these reporting requirements are based in part on the Know Your Customer regulations, which prohibit anonymous or pseudonymous bank accounts.

---

[16]For every cash transaction over $10,000, banks must file a Currency Transaction Report; for every $10,000 in cash or monetary instruments exported, the exporter must file a International Transportation of Currency or Monetary Instruments Report. Every business must report cash transactions over $10,000 by filing an IRS form 8300. This form is made available to federal tax investigators.

3.3.1.2 Motivation of Regulatory Information Requirements

In order to determine if a technological alternative, such as escrowed data or pseudonymous reports, can be optimized for electronic environments, the motivation for the reporting requirement must be considered.

It is neither reasonable nor necessary to go through every reporting requirement to assure that some alternatives are acceptable. Illustrating that changes can be made in certain reporting requirements for each possible set of techniques and motivation, is sufficient to illustrate that regulatory flexibility can result in auditing without surveillance.

Traditionally there are four basic reasons identified for the creation of reporting requirements:
- law enforcement
- tax collection
- optimization of social welfare
- risk management of the financial system (Heggestad, 1981).

Law enforcement in this case includes the Federal Bureau of Investigation, the Internal Revenue Service, the Drug Enforcement Administration and U.S. Customs Service. Data obtained by law enforcement using periodic reporting is made available through FinCEN to other agencies, including Interpol, the Postal Inspection Service and the Immigration Service (Office of Technology Assessment, 1995).

There is some correlation between the reason for a reporting requirement and a selected technique in that law enforcement cannot require periodic individual or aggregate reporting. This is partially a result of the Fourth Amendment: absent a warrant, we cannot be ordered individually to report our activities periodically to law enforcement. Data that is maintained for use in criminal investigations must be obtained with a warrant. Therefore the following examples of motivation and technique do not include examples for law enforcement in the periodic reporting categories.

3.3.2 Reporting Examples

In this section I construct a conceptual matrix, by providing examples in the sixteen possible combinations of motivations and techniques previously delineated.

Here I consider statutory reporting examples, not the regulations written to implement these laws. Regulations are more fluid than statutes, and thus less of a long-term concern. Regulation E provides an excellent example of this fluidity. Regulation E provides the specific implementation of the Electronic Funds Transfer Act. Regulation E has been frequently revised, most recently to better suit the capabilities of stored value cards (Federal Reserve Bank of New York, 1996). A similar rewrite of Regulation Z could alter the regulatory requirements of the Truth In Lending Act. Since regulations are bound by the statutes, easier to alter to meet current technological realities, and generally subject to change, currently I have focused here on the less tractable but more stable problems of statutory requirements.

3.3.2.1 Immediate Reporting

Consider an immediate reporting requirement for each of these categories: tax collection, optimization of social welfare, and risk management of the financial system. In the case of paper-based information systems, immediate reporting implies hours or days. Immediate reporting for the purposes of law enforcement as created in the Money Laundering Act was discussed previously.

An immediate reporting requirement at least partially for the purpose of tax collection is the requirement that exchanges of title to home be immediately reported. This allows property tax, charges for any building code violation, and other appropriate fines and taxes to be levied. This reporting requirement is strengthened by the fact that one cannot truly own a home until the title has changed hands in the public record.

An immediate requirement for the optimization of social welfare is the requirement that any officer of a company selling or buying stock in that company must report this transaction. This prevents officers from taking advantage of information about their own companies before it is released, therefore preventing manipulation of the stock market.

Immediate reporting requirements in risk management of financial systems do not exist per se. This is because most risk-seeking actions require approval in advance. However, close to an immediate reporting requirement is the requirement in the Truth in Lending Act that any changes in rates paid to customers be public and that banks not offer discriminatory rates. Not only does this prevent discriminatory pricing for social good, it also prevents banks from using discriminatory pricing to compete for the same few high return, frequently high risk, options. Most banks fulfill this requirement by mailing rate information to customers in their monthly statements.

3.3.2.3 Periodic Reporting

Now consider periodic reporting requirements for tax collection, optimization of social welfare, and risk management of the financial system. A periodic requirement for tax enforcement purposes, i.e. filing for tax payments or refunds, was previously discussed.

A periodic reporting requirement to assure the stability of the financial system is the requirement that all stock trades be reported to the Securities and Exchange Commission. The SEC cannot prevent actions such as insider trading and speculation that could weaken the market. However, full data on trades is necessary to detect insider trading, so that the criminal penalties can serve as a meaningful deterrent (Ziegler and Sanchez, 1993; Zuckerman, 1994). Furthermore, SEC regulations and detection of insider trading require not only the identities of those trading stock but also some attributes --- for example, employer and position in the organization --- to assure that senior executives are not taking advantage of privileged information.

Periodic reporting requirements for optimizing social welfare are contained in the Home Mortgage Act. The Home Mortgage Act requires that data on the specific mortgage requests which are accepted and rejected by banks be made available to the Federal Government. This provides an opportunity to identify and rectify discriminatory practices. The Home Mortgage Act (HMA) requires that the lending institution make a note of the applicant's age, gender and race on the application. The Home Mortgage Act does not then prevent the institution from keeping these records stored and linked to the applicant for future interactions, although the Equal Credit Opportunity Act forbids determining just these factors. This illustrates an opportunity for the ability of advanced techniques in information technology to remove this apparent conflict by requiring encrypted storage or highly-limited access to HMA records.

Separate from tax collection, there is no periodic aggregate reporting for the purposes of law enforcement. Reporting aggregate financial information to law enforcement would mean that groups which law enforcement has no reason to suspect and who have acted in no suspicious manner must periodically report to the police.

3.3.2.3 Periodic Aggregate Reporting

An example of periodic aggregate reporting for social equity as created in the Community Reinvestment Act was previously discussed.

Periodic aggregate reporting would appear to offer the least threat of data intrusion. In general, to assure that aggregate reporting does not become intrusive, the only requirement would be a limitation for microdata analysis for all but scholarly purposes. The reporting requirements that do create a threat to privacy do so by virtue of the data storage required for supporting documentation.

However the detailed data requirements necessary to make across-the-board reports can be intrusive. An example is the requirement that ethnicity and gender be reported for each issuance of credit included in the Home Mortgage Act. Thus without further consideration I move on to data storage requirements.

3.3.2.4 Data Storage

Consider an example of record keeping for each motivation: law enforcement, tax collection, optimization of social welfare, and risk management of the financial system. The Bank Secrecy Act is an example of data storage for law enforcement, as was discussed previously.

In tax collection, storage of all data relevant for purposes of taxation is required of any item or deduction that appears on a tax return until the period of limitation is over. For reported income or deductions, this is three years after filing or two years after paying, whichever is later. The period of limitation for unreported income is six years. If no return is filed, the Internal Revenue Service can demand documentation at any time. Thus the granularity of data storage requirements are controlled by the consumer.

The Truth in Lending Act was passed to prevent discriminatory and unsafe banking practices. The Truth in Lending Act requires that issuers of credit include in their reports to consumers the name of the merchant, the date and location of purchase. Furthermore, if there is some relationship between merchant and creditor such as a common parent or shared ownership of a subsidiary, then the item purchased must be reported as well.

To limit the exposure of the banking system, banks are required to keep of track of all outstanding loans and are not allowed to delete data from loans that fail. Interactions with directors and companies which have seats on a bank's board can be traced with this data. Also, individual votes by directors are recorded in order to detect, and hopefully avoid, conflicts of interests. This allows investigators to prevent a bank failure, or in the worst case, trace the cause after a failure.

These examples illustrate that a myriad of disclosure and reporting requirements serve a wide variety of purposes using the same set of technical requirements. Keeping this set of purposes and techniques in mind, consider the options in an electronic system.

3.3.3 Reconsidering Requirements

The previous set of alternatives offers interesting possibilities. I will revisit the set of examples and propose ways that the adoption of technical solutions can be encouraged.

Note that as electronic currency evolves not only are the possible methods different, the capabilities and desires of the market may differ as well. These requirements are:

- The government must have the data.
- The market has been unable to provide the data.

• There is no less intrusive reporting requirement under which the market could provide the data.
• The need is sufficient to justify the costs.

The costs to meet this need include the risks of decreased personal privacy, the cost to those required to report, and the administrative costs of the government. With this in mind, consider those cases where immediate reporting has been deemed necessary.

3.3.3.1 Immediate Reporting

Immediate reporting remains difficult to justify, yet can be necessary. It may be the least changed by electronic capabilities, with data transmission simply replacing the U.S. Postal Service. Yet the data reported may be less, since electronically reported data can be analyzed at the time of the report.

The options for reporting data include masking the identity of participants except as necessary for an investigation and using specialized software to analyze for suspicious activities.

Immediate data reporting is usually unnecessary. It is only when a particular case out of multiple data records is identified as worthy of investigation that the information becomes valuable. This suggests that constant pseudonyms could be provided for reporting on the activities of individuals.

Consider the promise of pseudonyms in the cases discussed previously. Immediate reporting is required when a home is purchased. This allows the local government to levy taxes and pursue violations of building codes. What information needs to be accessible on-line in order to provide easy access for those with legitimate need while not providing easy access to the price of your home and the sum of your holdings? This is a case where limits to disclosure or conditional psuedonymity apply. If a home owner pays the bills and maintains his building up to code, then there is no reason for anyone, including marketers, to easily access information about the value of a home and its owners. Listings of buildings which owners do not maintain at code would allow community groups and others with an interest in contacting the owner of a specific property to identify that person. Similar arguments hold for automobile ownership records.

Conversely, the identity of the individual could be stored with separate agencies so that cooperation is required to link an identity with a purchase. This can be done using secret sharing (Shamir, 1979). This would allow for searches under names when an individual is charged with a crime. This would allow identities to be made available under special circumstances while preventing the widespread dissemination of information to marketers for further analysis.

Currently, the only immediate reporting requirements for financial transactions are based on a threshold, i.e. size of the transaction. These reporting techniques can be made less intrusive since electronically reported data can be immediately analyzed for suspicious activity. For information where there is no suspicious activity, the data can be deleted by the receiving agency. Thus, it is possible that while less information is directly reported more suspicious activity is identified. Reporting many transactions with pseudonymous identifiers could make it harder for individuals to smurf[17], while decreasing overall intrusion.

---

[17]To “smurf” is to divide one large transaction into multiple small transactions to avoid reporting requirements and subsequent detection.

The purchase of stock by a director is a rare case where the identity of the person is important because the position of the person must also be identified. A stock purchase by an individual in charge of mergers and acquisitions is more interesting than a sale by counsel when no litigation is pending. Here individuals limit their privacy by choosing a position of responsibility where a higher level of oversight is necessary. A similar argument could be made in the examination of financial decisions made by legislators and powerful members of the executive branch.

The reporting required in the Truth in Lending Act is an excellent model. Banks are not required to report the identity of customers to the state to assure that all customers have been notified. The banks need only show that a policy exists and procedures are followed.

3.3.3.2 Periodic Reporting

Consider the cases of periodic reporting previously discussed. The Home Mortgage Act assures that individuals are not discriminated against. This differs from the Community Reinvestment Act in that the Community Reinvestment Act was passed to prevent targeting of communities. The reporting requirements in the Community Reinvestment Act and the Home Mortgage Act can be unified with loans allowed or identified by ethnic origin of requester, ZIP code, minimal financial data, and sex. Financial institutions could then be required to avoid the statistical appearance of discrimination.

In order to avoid fraud in reporting, a cut-and-choose technique could be used to statistically identify fraud. The number of records checked increases the certainty that there is no fraud, but eliminates the privacy of individuals whose records are so chosen. Mortgage information is already on-line, and is in fact sorted so that information can be sold to other providers of financial services (Fenner, 1993). Thus, using a cut-and-choose technique is not as costly a proposal as it might appear.

Electronic stock reporting may be simplified by reporting the number of shares bought and sold with a conditional pseudonym. Insider trading can be identified by investigating suspicious transactions; then only those pseudonyms are exposed. Since the total shares of stocks purchased and sold can be determined by watching the market, the existence of unreported sales can be identified by statistical techniques. The identity of the individual in questionable purchases could then be obtained from the broker.

3.3.3.3 Periodic Aggregate Reporting

Periodic aggregate reporting is problematic in that it implies storage of individual consumer data in order to obtain the aggregate data.

If consumers had smart cards then the prevision of periodic reporting without identity information could be possible with anonymous updates of aggregate information (Camp and Tygar, 1994). The creation of smart cards offers one solution for this; however, individuals will inevitably lose such cards. This implies that those with unattractive data could simply lose the information. Mandatory back-ups at a trusted facility could mitigate this problem.

3.3.3.4 Data Storage

A fundamental problem with reporting data is that it requires that this data be captured. Once data has been captured there are problems with internal disclosure, external disclosure and the resulting creation of privacy-threatening compilations.

Current limitations on disclosure apply only to the Federal Government. Furthermore, financial data has limited protection under the Fourth Amendment, as clearly stated in

United States v Miller. Thus, limits on disclosure of information to government are weak, but they do exist.

Internal disclosure occurs when required data compilations are used within the financial services institution. When the investment in data collection is required by financial regulation, allowing institutions to use data internally softens the financial blow.

With concentration or monopolistic powers, the use of such data can become more problematic. As transactional information becomes more detailed, the use of this information in internal hiring, promotion, and consumer credit provision decisions becomes an increasingly important issue. Can a bank look through consumer records to evaluate applicants? Currently the bank has the right to do so, since it is valid internal use.

What of internal disclosure for purposes of prosecution? An example which is not too extreme is the possibility that Microsoft could use information obtained through its network services to identify possible violators of electronic copyright. There is currently no prohibition against Microsoft using its own internal data to assist law enforcement in identifying possible thefts of software. Legally, it does not matter that the data was obtained without the consumer's knowledge.

Consider now external disclosure. Any data that can be used for internal decisions may be sold to other organizations for similar decision-making purposes. There are currently no constraints on the commercial trade of such data. In fact, once the government has obtained data, it must release it to requesting organizations regardless of their motivation.

In light of the motivations and the options created in electronic information systems, information storage requirements seem unnecessarily intrusive. The Truth in Lending Act storage requirement could be changed so that the customer need only have a valid signed agreement, and upon presenting that receipt can obtain a full refund. Current electronic commerce protocols can be designed so that storage by the credit-grantor of items bought is not necessary for the provision of receipts -- encrypted signed receipts or transfer of purchase orders provide nonrepudiation. The practical requirement that information about purchases be recorded in easily searched, correlated, and reproduced format exacerbates the threat of data surveillance.

Before requirements for transactional data are created, the threat of possible surveillance should be balanced against the wrong being addressed. This suggests that any compilation of data about consumers by other parties which is required by law should also be protected by law.

3.3.4 Summary

In this chapter I have illustrated that current policies reflect an awareness of the utility of electronic information, rather than an understanding of the threats of emerging technology. Secure, reliable and private commerce systems are hindered in their development by the current combination of widespread availability of true anonymity and a legal regime which arguably promotes privacy violations.

The response by the legal and regulatory communities to privacy-threatening innovations, and to new technologies in general, has been the development of technology-specific rulings after these technologies have been dispersed through the marketplace. This approach is breaking down as the rate of technical change increases. It may not be possible to respond to information technologies after they have reached some critical mass; privacy protections may need to be included in the hardware (Morgan, 1992). Arguably in this information age there needs to exist a system of laws recognizing that the right to privacy is technology independent.

In the conflict between privacy and data reporting, many problems associated with paper currency remain with electronic currency. Among these common problems are law enforcement concerns, tax collection, auditing and fraud detection, prevention of discrimination in the provision of financial services, financial privacy, the assurance of funds for socially desirable goals, and the balance between risk-reducing regulation and productive but risk-seeking free market behavior that best moderates banking cycles for the least cost. There is no reason to abandon the goal of solving these problems as currency becomes increasingly electronic. But as the nature of currency and commerce change, previously chosen methods of advancing regulatory interests are increasingly ill-suited. Just as different techniques are appropriate for paper currency, different techniques are appropriate for electronic applications.

This is particularly difficult in the case of transactional data. From the perspective of customers, information about their purchases is clearly personal information. For merchants and electronic commerce providers, it is critical business information about consumer preference, as well as a product in its own right.

The current debate over customer information in the increasingly competitive voice telephony market may provide a glimpse of the conflicts to come. Customer Proprietary Information is data about whom a customer calls, and when. This information has tremendous privacy implications. It also has increasing market value, especially as local exchange markets become competitive. Knowledge of the calling patterns of a region, neighborhood, or an individual is a powerful weapon in the competitive marketplace. In order to provide a level playing field all possible entrants into a communications market should have the same information. However, this implies that all information about the location and duration of every phone call a consumer makes should be available to anyone who can claim a possible competitive interest.

Customer Proprietary Information is both important commercial and private personal information. There is an expectation of privacy in the contents of one's phone calls, and a tradition of privacy in the Bell System. Yet there is also both a public interest in fair competition that would compel the release of such information. This type of conflict will become increasingly common as information technology proliferates. The solution to the CPI debate -- widespread information availability -- is not entirely promising. Yet the intelligence in the hardware at the endpoints of Internet commerce offer a broader range of possibilities than is the case in telephony.

One clear conclusion in this examination is that laws written for rapidly changing areas, such as funds transfer and consumer leading, should be written with the least technologically restrictive language possible, so that innovative solutions may be allowed or required by regulators without waiting for an act of Congress.

4 Separation & Examination of Internet Commerce Systems

“The greatest pleasure of ignorance is the pleasure of asking questions.”
Parker, 1959

Reliability, security and privacy are important in commerce systems. Current commerce systems include examples as diverse as paper cash and Internet commerce and hardware-based commerce. Credit card verification systems, point of sale transfers, lines of credit, billing servers, secure co-processors, and systems based entirely on software all co-exist and compete for consumer and market interest. There are also special purpose systems such as electronic postal metering systems and copyright collecting services.

Electronic currency systems are as widespread as they are diverse. In the 1970's electronic currency systems such as electronic funds transfer (EFT) began to be widely used (Reid and Madam, 1989). Currently the net value of all electronic transfers exceeds the total value of all cash used. In 1990 over 40% of the $500 billion in federal benefits and state-administered programs were paid using some form of electronic funds transfer (Wood and Smith, 1991). In 1988 physical rather than electronic currency, including cash, credit cards, or financial instruments, accounted for over 99% of all transactions (Newberg 1989). The same statistics reflect transaction patterns today; the total value of all electronic transactions dominates the total value of the vastly more numerous cash transactions.

In addition to the widely used ATM systems and the Fedwire, there are private networks and products which provide automatic budgeting, check writing, and invoice creation. The sheer number of electronic currency and invoice systems available for private institutions overwhelms the possibility of an exhaustive report.

Given that an exhaustive report is not feasible, which systems should be considered? Separating systems into categories removes the need for detailed analysis of every system without unreasonably limiting the scope of this work.

Recall that I am concerned with systems for general Internet commerce. These systems require that the user have no specialized hardware.

4.1 Separation

First, systems are separated into *token* and *notational* systems. In token currency the strings of bits transferred in a transaction are themselves legitimately valuable. For example, a dollar has value in and of itself, and is not a promissory note for a particular transaction from a specific account like a credit card purchase slip. In a notational currency system the information transferred is an instruction to change notations in a ledger, such as a bank's records. In notational currency the value is held in the records, not the instruction.

Second, notational systems are separated by business model. The business model includes the underlying commerce model, the distribution of risk, the distribution of liability, and therefore the distribution of trust.

Notational money exists as notations in the ledgers of an institution. Electronic commerce systems based on notational currency are separated by the role of the institution which holds the ledgers. Customers and merchants must trust, to some degree, the holder of the ledgers.

Notational systems are based on different models, where customer and merchants pay different fees and take different risks. Notational systems may be based on the checking, debit card or the credit card model. In some systems there are financial intermediaries which alter the traditional assumptions of interaction between merchants, customers, and their respective financial institutions. The existence and roles of these intermediaries may change the distribution of liability, and therefore trust, in financial transactions.

The implementation of anonymity in token systems underlies trust requirements. This is because of the relationship between anonymity and accountability: an anonymous party cannot be subject to penalties. Anonymity in token systems determines who is able to take untraceable, and therefore potentially deniable, financial action.

Other characteristics may be incidental to a system and can be changed in a given implementation. This fact allows me to collapse the systems into classes within the framework previously developed. Then, within each class one system will be selected as an example for more detailed analysis.

The token systems that I analyze and their respective levels of anonymity are:
• Digicash: token, complete anonymity
• Digicash with embedded user information: token, conditional anonymity
• Micromint: token, no anonymity

The notational systems I will analyze are:
• First Virtual: bank off-line, transactions without security
• Secure Sockets Layer: bank off-line, secure transactions
• NetBill: single bank on-line
• Secure Electronic Technology Specifications: multiple acquirers with on-line presence and mutually respected certificates
• Anonymous Credit Cards: customer, merchant and intermediary banks have on-line presence and are separated for reasons of privacy

4.2 Analysis

After selecting examples from each technical category, I will analyze these systems along legal, market and technological axes. The focus of the discussion will be the ability of the systems to meet market and legal constraints.

4.2.1 Transactional Reliability

I begin by discussing the perspective of each plan. If a protocol is modeled on a specific physical currency, I note this.

In this section I say which parties are considered trustworthy, the bank, for example. The business perspective may explain how a technical failure is not unacceptable, since the realization that such failures occur is addressed explicitly in the business plan.

Both reliability and security depend on the business plan. If the business plan is flawed, a lucrative security hole may exist as a result. Thus security is not simple algorithm choice and key management but also an understanding of the business implications of any system failures.

Similarly, atomicity depends on design, implementation and business policy. Atomicity depends on funds available policies because of rollback, as explained in Section 2.2.2.

For one protocol in each analysis chapter I specify the fields and cryptographic operations in the protocol. There are four functions commonly used in electronic commerce protocols. These are a one way secure hash algorithm, asymmetric encryption and symmetric encryption. The following notation is used, when the variable or message being encrypted or hashed is *x*.

h(x)	the hash of x
$E_k(x)$	x, encrypted with symmetric key k
$(x)_i$	x, encrypted with the secret key of i's asymmetric keys
$(x)_I$	x, encrypted with the public key of i's asymmetric keys

The public and private halves of a public key pair are identified as I and i, where i is the first initial of the party or item to which a key corresponds. For example, values of I might include B, C, and M; these correspond, respectively, to a bank, the customer, and the merchant.

I consider each transaction step by step. The focus of this analysis is the determination of the reliability of a transaction. In a step by step analysis, I can illustrate how the failure of a single message might put the system in an inconsistent state. Each protocol is classified as providing no atomicity, money atomicity, goods atomicity or certified delivery. (Classes of atomicity were defined in Chapter 2.) As part of such a classification the specific messages which provide or fail to provide atomicity are identified.

4.2.2 Security

After a reliability analysis, I consider the possible failures of explicit security assumptions. Every system has certain security assumptions, whether it is the existence of separate communications channels or the assumption that secret keys are indeed secure.

For each system I determine the worst case results of the failure of each security assumption. I do not consider the possibility of every combination of security failures. I do consider the case where the simultaneous failure of two security assumptions creates a different possible outcome than the separate failure of two security assumptions. For example, if both a secret key and an account number would enable an attacker to make unauthorized changes in an account, then I note this fact. However, when joint failures are required to enable fraud I address joint failures. For example, I consider the case where the loss of a bank’s secret key and a consumer’s secret key would enable the attacker to, respectively, counterfeit cash and withdraw it in a customer’s account only if such fraud would not be possible from possession of only one set of keys.

I do not consider the failure of certain fundamental, widely accepted cryptographic assumptions. For example, though it has not been proven that there is no way to factor numbers in polynomial time, I will make the standard cryptographic assumption that this is the case (Baker, 1984; Schneier, 1995).

4.2.3 Privacy

The next step in my analysis is to evaluate the privacy considered in each system. To do so, I list the information available to each party during each transaction in a matrix. This allows for simple comparison of different systems without trying to place a qualitative measure on the intrinsically quantitative issue of privacy. There is broad agreement that privacy is a question of what information is exposed. Thus I focus on information exposure rather than strict classification. An example is shown below.

To consider the efficacy of this method, recall the specific methodology for evaluating system security. There are elements of this methodology applicable to the analysis of privacy levels in a system. The potentially applicable steps include: a general description of all information that will be transmitted through and stored in the system; a summary of the expectation of the security of data contained in each subsystem and system; the assignment of final responsibility to a single individual to assure security is maintained; the use of mechanisms to assure security; a description of the entire user community including those with the lowest level of access; and the types of access permitted provided in each subsystem and system. The simple matrix technique used for assessing privacy in the following system analyses includes many of these elements.

An example matrix, which shows the information available to various parties in a checking transaction, is shown below in Table 4.1. Each row shows the information available to the party named in the leftmost column. Each column lists a datum of interest: the identities of the parties, the date of the transaction, the amount of the purchase and the item purchased.

Information Party	Seller	Buyer	Date	Amount	Item
Seller	Full	Full	Full	Full	Full
Buyer	Full	Full	Full	Full	Full
Law Enforcement w/ warrant	Full	Full	Full	*Full*	None
Bank	Full	Full	Full	Full	None
Observer	Full	Full	Full	Full	Full

Table 4.1: Information Available to the Parties In a Check Transaction

The row labeled "Law Enforcement with warrant" identifies the information that would be available to the government. This row provides a basis on which to compare each system's ability to provide societally desirable information. (Socially desirable information is characterized in Chapter 3.) Law enforcement can obtain records from bank and merchants. However, in the case of a specific purchase, the merchant may not keep detailed records. Here and in later tables if Law Enforcement depends on merchant records to obtain item information, then it is denoted through italics as *Full* .

For the privacy analysis, we consider an observer who is electronically well-placed, i. e., the observer can monitor transmissions between the customer and merchant. The observer cannot read encrypted information but can read all other information.

Using an information matrix as above as a basis for comparison, each system is roughly classified as providing high, low or medium levels of privacy. Any party can have full, partial or no information about the datum of interest.

I focus specifically on the information made available in the protocol. This provides a consistent comparison of the protocols. However, it is important to realize that information transmitted during the discovery process is available with every purchase. Thus those protocols that claim no identity information is transmitted in a transaction assume a communications exchange or remailer. Removing all identity information during communication is not trivial. For two-way communications this requires a series of remailers capable of encryption. Even then partial identity information can be reconstructed with the cooperation of all involved parties; however such a reconstruction has such a high work factor that it is reasonable to say such services provide privacy.

The reader should keep in mind while reading the privacy analysis of each system that partial or full identity information can be obtained from IP addresses and commonly available network services. Recall that this was discussed in Section 2.3.2. I will compare the Internet protocols only on the basis of information made available after discovery. This provide a more clear differentiation of protocols.

4.2.4 Regulatory Issues

Finally, I consider the ability of the system to fulfill governmental needs for data. I consider only the information provided about a transaction itself. It is true that most transfers into traditional currency would use audible channels. The general policy implication of this is that anonymity constraints should be relaxed.

In Chapter 3 I both identified reporting requirements and offered suggestions for improving policies requiring information availability. In the regulatory fit section of each analysis, I revisit these specific suggestions for changes in law. In each case I identify the trade-off created by the current requirement and the way that my proposed solution would alleviate it.

In the regulatory fit section for each system analyzed, I also make technical suggestions on how the system could be enhanced to resolve or eliminate trade-offs. Techniques used in this section are described or referenced in Chapter 2.

5 Notational Currency

“In a sense a person is defined by the checks he writes. . . . The banking transactions of an individual give a fairly accurate account of his religion, ideology, opinions, and interests.”
Justice Douglas, 1974

“The common law secures to each individual the right of determining, ordinarily, to what extent his thoughts, sentiments, and emotions shall be communicated by others. . . . The existence of this right does not depend upon the particular method of expression adopted.”
Justices Warren & Brandeis, 1890

In notational currency, the information transferred consists of instructions for payment. The value in this currency is stored as notations in the ledger of a trusted institution. Transactions require that these notations be changed.

The advantage to notational currency is that record keeping is an inherent part of the system. This simplifies recovery from failure. If a single ledger is used the transaction is certain to be serialized and as a result implementing ACID transactions is straightforward.

Here I analyze five Internet commerce systems which use notational currency: First Virtual, NetBill, Anonymous Credit Card, Netscape's Secure Sockets Layer, and Mastercard's Secure Electronic Transactions. I use the approach described in Chapter 4. I also examine current telephone orders using credit cards. Remote notational currency transactions are already common in the form of telephone orders. Many of the problems with mail-order and telephone-order purchases exist in Internet commerce. Thus I begin this chapter with a discussion of traditional telephone order transactions. This provides a clear introduction to the complex transactional issues in notational currency.

5.1 Credit Cards

With credit and debit cards, the instruction to debit or increment an account is made electronically. When a purchase is made over the telephone, the information printed on the card is sufficient to authorize a charge. Thus credit card orders are authorized using *information only*. This differs from point of sale (POS) systems, in that the physical presence of the card is not necessary in a telephone order.

Remote credit card transactions are a form of electronic commerce; the critical transaction information is delivered electronically by voice, from one human to another. Orders are entered into billing computers, processed electronically, and delivered physically.

The credit card market developed because of the limits of interoperability in checks (Rubin and Cooter, 1994). This interoperability was a result of the fact that merchants were at risk for insufficient funds checks and therefore accepting a check required an extension of trust to the consumer. Previously customers were limited in accessing funds because merchants had to extend trust to every customer who wrote a check. The creators of credit cards, or entertainment cards as they were originally called, addressed that weakness by assuming the risk that customers were not credit-worthy.

Today automatic teller machines now offer international interoperability between checking accounts by providing immediate access to deposits in the form of cash. That the original impetus for the creation of credit cards is gone does not mean that the credit market will decrease, since the original entertainment cards have evolved beyond their original market niche.

The settlement process for credit cards is shown in Figure 5.1. While this is the specific system for Visa, it is representative. The credit settlement system is similar to the check settlement system. However, credit cards developed in a more orderly fashion than checks, and have evolved a clearance system with less need for governmental support.

Notice that the credit card system is based on regional hierarchies. Each region has a clearing bank. This means that Visa does not have to handle every charge. Each clearing bank has a network of local banks which recruit merchants. Card-issuing banks may further subcontract credit checks for individual transactions to third parties.

When a bank recruits a merchant it accepts some risk. If a merchant defrauds a customer, or a customer stops payment on the basis of fraud, the merchant bank is liable. The merchant is guaranteed payment from the merchant bank. This liability is the control mechanism used to prevent unethical merchants from entering the system and taking advantage of the assured payment mechanism. It does this by penalizing the banks, which control entry into the system.

Payment is assured by distributing the loss across all purchases in the form of fees. Each merchant receives a percentage of each purchase. This is illustrated by the decrements of funds shown returning to the merchant in the outer loop in Figure 5.1.

When a customer is defrauded, the customer can refuse charges. Whether this is a result of physical theft or electronic attack, the customer's losses are limited by law to $50 per card.

Customers have sixty days to refuse a charge, much longer than with a check. Therefore customers do not feel the need to trust merchants as much as in the case of checks. Thus customers expect less documentation. This allows credit charge slips to be radically truncated, that is the merchant sends the data from the slip, not the slip itself, to the bank. While it is legal for checks to be radically truncated, customers have demanded that their checks be returned.

5.1.1 A Transaction

Figure 5.1 shows the steps in a credit card transaction.

The steps in the transaction form two concentric semi-circles: the billing records for the customer travel in the center counter-clockwise circle and payment to the merchant is represented by the arrows forming the outer, clockwise circle.

In Figure 5.1 the inner circle is the movement of the order information; the outer circle shows the movement of payment in return. The merchant pays a percentage to the merchant bank on every charge. Depending on the specific authorization and the contract with the merchant bank, the merchant may be guaranteed payment. However, in remote credit card orders it is usually the case that the merchant will not be paid if the customer denies the charge.

The difference in the amount paid by the customer for the purchase and the amount received by the merchant goes to profit, overhead and risk management. The lack of atomicity in credit card purchases is managed through this systemic debiting as the payment instructions move through the system. Notice that when the merchant and customer banks are the same, Visa does not receive the transactions fees.

A credit card transaction is not money atomic, although it can appear atomic to the merchant if the merchant is guaranteed payment. From the customer's perspective, credit card purchases have a period in which payment can be canceled, either by explicit cancellation request or by a refusal to pay for an item when billed.

Credit card transactions are consistent in that the customer and merchant agree on the amount paid and whether there was a reversal of payment.

Credit card transactions are not isolated. There are cases in which a merchant obtains a block on user credit which is not always promptly erased. These can in some cases lead to failure of isolation. For example, a hotel may block enough to cover possible damage, thus preventing the customer from making a later, unrelated purchase.

Credit card transactions are durable. However, it may take weeks for a credit transaction to become final.

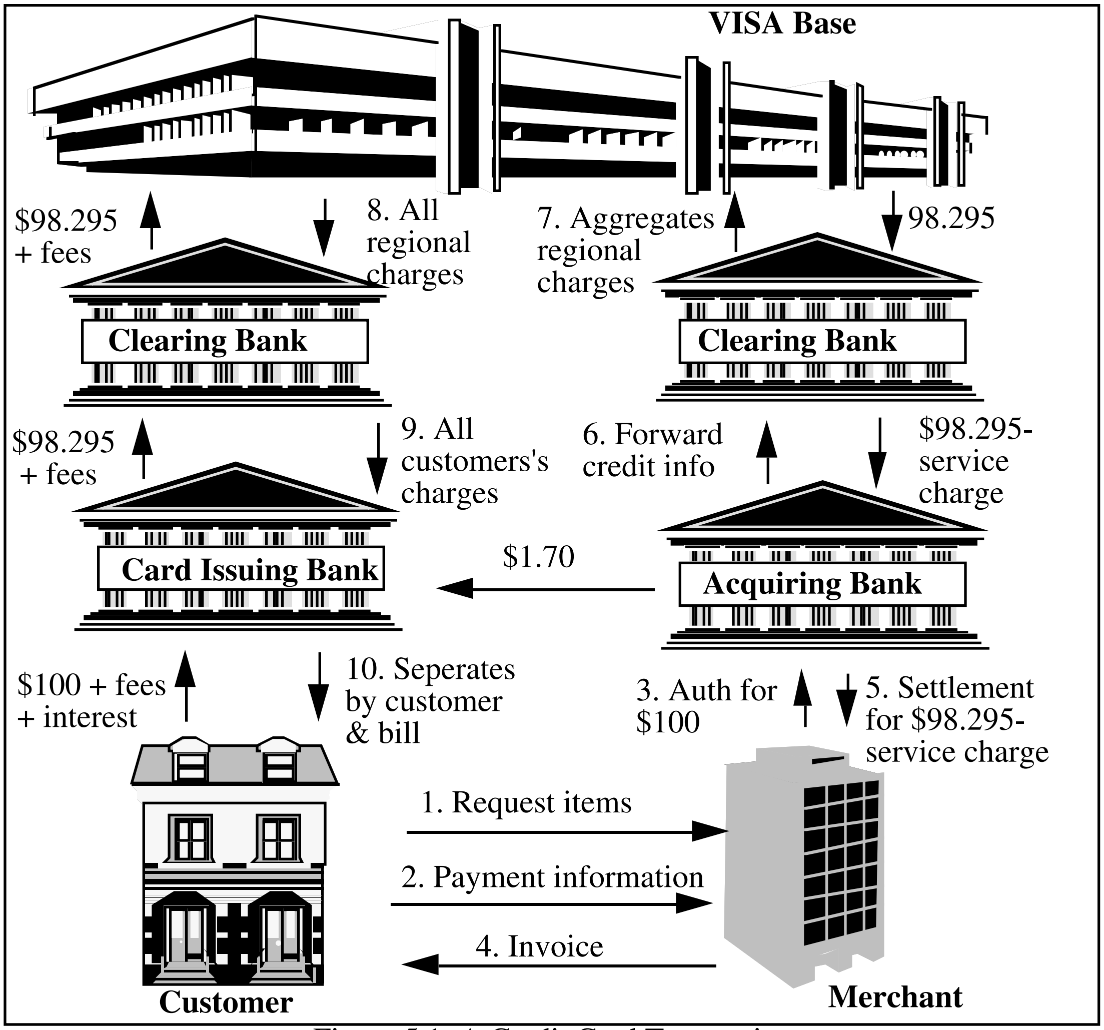

Figure 5.1: A Credit Card Transaction

A credit card transaction is limited only by the customer's credit limit. With ATM machines, there is a size limit for an individual transaction. This is to limit risk, just as the limit on currency denominations limits risk by increasing the difficulty of counterfeiting -- in both cases the limits increase the number of false payment instruments needed for large scale fraud. However, this sometimes fails, as the recent theft of over $300,000 with a single card and access code illustrates (New York Times, 1995a; New York Times, 1995b). Like checks, credit cards can support a nearly limitless number of users.

There are limits to interoperability between credit and debit card systems. A consumer cannot pay her American Express card bill with her Visa, except by first obtaining cash. (This statement has been experimentally verified by the author.)

5.1.2 Security

Opportunities for fraud vary with payment policies. With remote credit card purchases there are two options: the acquiring bank guarantees payment or the merchant accepts risk.

In the case that the bank pays, merchant fraud is trivial. Merchants can fail to deliver goods and demand payment. Banks limit this fraud by tracking complaints against merchants and revoking merchant accounts. Merchants can re-incorporate and request new accounts with these new identities. However, some merchants allow other disbarred merchants to use their accounts, so this system itself has weaknesses (Van Natta, 1995).

In both cases customer fraud is straight-forward. Customers may simply claim not to have received goods. The lack of a verifiable physical delivery system constrains security for all remote purchases of physical goods. Banks address this in the same way that they address merchant fraud: customers are tracked, rated, and subject to the removal of credit privileges.

If the merchant has a physical presence, i.e. an imprint of the card, the merchant is guaranteed payment. If the merchant is offering telephone purchases, the certainty of payment depends on the type of merchant, merchant characteristics and transactional characteristics. Important merchant characteristics include credit history and market served. Important transactional characteristics include price and item purchased.

Arguably the greatest weakness in the telephone order protocol is its vulnerability to replay attacks. Only credit card number, expiration date, and sometimes billing address are needed to authorize a purchase. Thus, any person who obtains a complete receipt can authorize telephone purchases.

This weakness is exacerbated by the fact that this information must be transmitted over the public telephone networks. However, the public phone networks are more difficult to monitor than Internet transactions for several reasons. First, the telephone network is built to transmit voice data. Filtering and searching voice data after it has been trapped requires decoding the data and then using voice recognition technology to obtain the content. Voice recognition is orders of magnitude more difficult than identifying information already optimized for digital content-based manipulation. After having analyzed the data to determine the content using voice recognition, then the data must be searched, just as in the case with data on the Internet. Conversely, information on the Internet is typically in a form that is simple to intercept and filter. Finally, the sheer scale of the two systems makes monitoring phone calls harder. There are 28 million (Hoffman, Kalsbeek and Novak, 1996) to 37 million (CommerceNet, 1995) Internet users in the United States, while there were 172.8[18] million households (Bureau of Census, 1995) subscribing to telephony services in 1995. Telephone service is common to almost all households, with 96% of households having telephone service in 1994 (Federal Communications Commission, 1995).

5.1.3 Privacy

Credit card transactions create machine-readable records. Identities of the parties, price, date, and some content information may also be recorded with credit card purchases. Privacy can be compromised by the ease with which this information is analyzed and distributed.

---

[18]This includes both access lines and cellular subscribers.

Credit card purchases provide detailed information about the transaction to merchants and associated financial institutions. They can also leak information through electronic surveillance by an observer. Because information is obtained by the bank, it can be available to law enforcement. Content information may be recorded by the card issuer, and content information in machine-readable format may be trivially obtained by the merchant. Content information is not typically transferred by merchant to acquirer. Information distribution in a remote credit card transaction is enumerated in Table 5.1.

Note that I combine the merchant bank with the clearinghouse, the customer bank and the intermediate banks. This implies an assumption that these institutions are all one bank -- otherwise the customer bank would be the card issuing bank and the merchant bank the acquiring bank. This assumption is supported by the fact that there is widespread marketing and sharing of data between banks.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	Full	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf w/warrant	Full	Full	Full	Full	*Full*
Bank	Full	Full	Full	Full	None
Observer	None	None	Full	None	None

Table 5.1: Information Available in a Credit Card Transaction

Telephone transactions require sending information over phone lines. Recall voice-based telephone transmissions are not as easily intercepted or as subject to electronic monitoring as Internet transmissions. However, if such a transaction is intercepted then the observer has all the information necessary to charge orders to the customer's account.

5.1.4 Regulatory Issues

Since credit cards are now widely used, many regulatory issues have already been considered. In fact, two specific regulations of interest were enacted at least partially with consideration of credit cards: the Electronic Funds Transfer Act and the financial information provision of the Computer Fraud and Abuse Act. Thus a short consideration of lessons learned with credit cards may be fruitful for later discussions.

The Electronic Funds Transfer Act was passed because of the recognition that a customer has neither the ability to strengthen the payment system nor the ability to prevent use of a financial instrument once stolen. The initial assumption that a customer would bear these charges in the event of a lost card is reflected in the assumption of many electronic commerce systems that the customer will simply bear lost charges in the event of a lost cryptographic key. In fact, under the Digital Signature Law passed in Utah, if an attacker obtained a consumer's secret key, the attacker could enter in contracts requiring the customer to continue to pay for many years. This results from the lack of consideration of key loss -- keys are treated as if they were unalterably linked to an individual, like the signature on which keys are modeled in this law.

In terms of privacy, credit card companies policies vary widely. American Express offers consumers options on the use of personal information. Mastercard and Visa allow card issuing banks to set policies on consumer privacy. That one company offers a privacy-protecting option suggests that the market can serve the privacy interests of those with financial resources.

5.2 First Virtual

First Virtual is based on the theory that the provision of information goods over the Internet is practically free and that the Internet itself is inherently without security. First Virtual aggregates transactions, filters transactions, provides billing, and resolves disputes (First Virtual, 1995a). First Virtual is an account acquirer from the perspective of consumers and merchants; in contrast, First Virtual is a single merchant from the perspective of the acquiring bank.

First Virtual filters transactions and resolves billing disputes about these transactions. First Virtual resolves disputes by maintaining that the customer is always right. First Virtual limits customers' abuse of this policy by limiting the customer's total number of refusals -- after a given number of refusals a customer's account privileges are terminated.

First Virtual's protection against fraud is based on three business practices:
• no credit card numbers are ever on the Internet
• no replay attacks are possible
• the losses of a merchant who is unpaid for network-delivered information goods are negligible

First Virtual is a protocol for the first generation of Internet commerce. As with all on-line systems First Virtual has automated customer support, promotion, administration and processing. Some First Virtual transactions are large enough that aggregation is unnecessary; however, First Virtual can provide aggregator services for small transactions. The goal of First Virtual is not to decrease transaction costs but rather to provide immediate access to customers on the Internet for medium-priced information goods.

Commerce without security has limited application. The size of a purchase with First Virtual is limited by the merchant's tolerance for fraud. Merchants with high cost goods for which there is a high demand are unlikely to accept high levels of fraud. First Virtual works well for low priced goods with a small to medium market, or high priced goods with a specialized market. The acceptance of First Virtual is also subject to the frequency of attacks on customer account identifiers and customer tolerance for the time and effort in addressing these attacks. If First Virtual claims a charge is valid and the customer disagrees, the customer still has the protections against loss and fraud provided through her card issuing bank.

First Virtual is useful for information goods delivered over the Internet. The market for on-line information goods is hampered by the fact that goods are widely distributed and often have very low value. Many on-line information merchants are not large enough to have merchant accounts with credit card companies. Besides the number and small size of many information providers, this is a problematic market for current Internet commerce protocols because the value of these merchants' items is so low, consumption happens soon after delivery, there is no standard for proof of delivery on-line, and there is no physical presence. Because First Virtual's business model is based on negligible merchant losses, First Virtual is not well suited for orders for physical goods.

Becoming a First Virtual merchant requires a credit card, email, data storage capacities and Web access. Notice that an Internet user can be a First Virtual merchant with a standard bank account, while other systems require that merchants have merchant accounts. First Virtual's approach vastly expands the number of possible merchants, and therefore the probability that there will be information of interest to a customer.

Obtaining a First Virtual account requires email and a credit card. The prospective customer sends email to First Virtual that includes a customer-selected password. The customer then calls First Virtual and provides credit card information over the telephone. The credit card information itself is never sent over the Internet. The password and user name (which is called an *account identifier* by First Virtual) are used by customers to authorize charges against their accounts. There is a low initial fee to become a First Virtual customer.

Presumably, First Virtual also profits from the redistribution of the email addresses of its customers. Customers are allowed to opt out of this program when asking for an account; however, the default value is redistribution.

5.2.1 A Transaction

A First Virtual transaction is shown in Figure 5.2. Note that the bank is actually off-line, and is contacted by First Virtual after a transaction or a series of transactions have been completed.

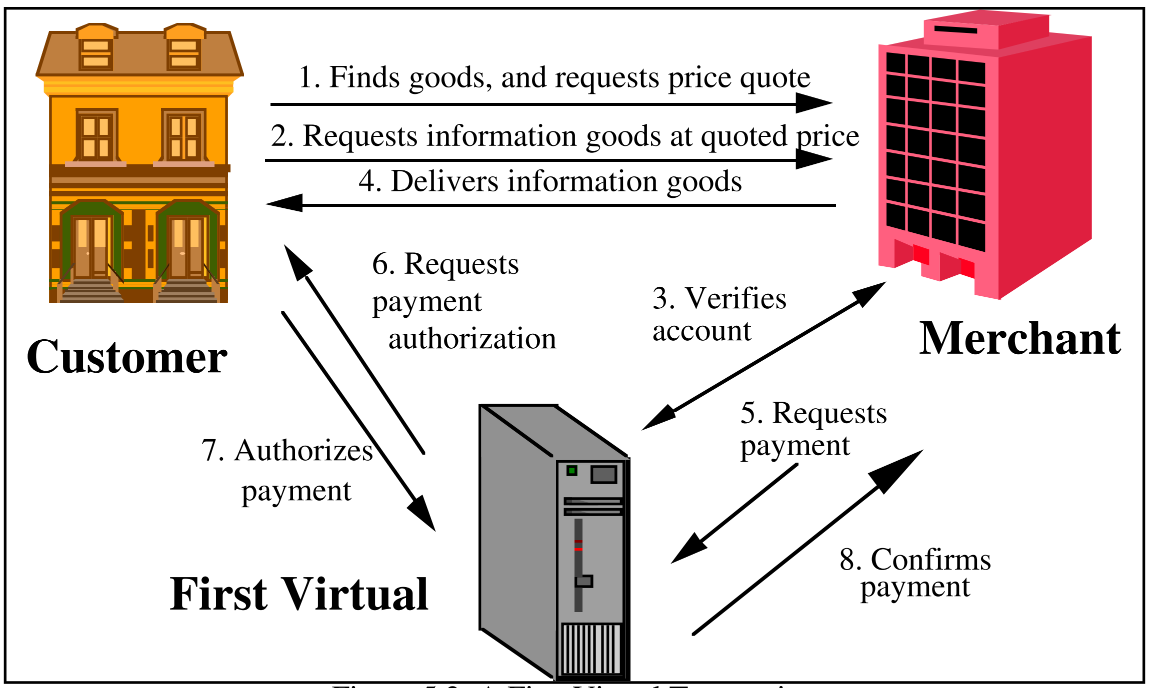

Figure 5.2: A First Virtual Transaction

The transaction begins when a customer selects an item. The customer then requests the item with a message that includes her First Virtual account identifier and the associated password. The merchant could authenticate the customer's claims to be a valid First Virtual customer at this point, as shown in step three of Figure 5.2. First Virtual will verify that password and account identifier. If the customer is a valid First Virtual customer, the merchant is then contractually obligated to deliver the requested items.

After the merchant sends the goods the merchant sends the customer's payment authorization to First Virtual and requests payment, as shown in step five. The merchant is required to send the merchandise requested before asking for payment. First Virtual then sends an email message to the customer for final authorization of the charge. The customer will only be charged if the customer verifies the charges. Finally, First Virtual notifies the merchant of the result.

The merchant may choose to validate the customer at step two, three or five. Validation at step two means that the merchant never accepts a request with an invalid password. There is a trade-off in this choice: a merchant saves one message on every valid transaction or saves processing invalid requests. The appropriate choice depends on the ratio of valid purchases to fraudulent requests as well as the relative costs of communications and processing.

Regardless of the merchant's timing of verification, the customer has the right to refuse to pay for an item after having received it. This prevents conflicts based on quality and deceptive advertising. First Virtual reserves the right to limit the number of times a consumer may choose not to purchase an item received; but a merchant can not choose to refuse to send an item to a valid First Virtual customer. This means the merchant must accept First Virtual’s definition of acceptable risk.

The email in step six and the request in step two travel through different parts of the Internet, like a telephone call to Tokyo and a fax to New York travel through different parts of the telephone network. Therefore First Virtual considers these independent channels. While it is simple to obtain a packet containing ordering information from First Virtual, intercepting the authorization request message to the customer is more difficult. It would require either filtering every message received by the customer or sent by First Virtual, or alternatively breaking into the customer’s home email account. Furthermore, there is no gain in completing the second, more difficult, part of the process because any attacker has already obtained the goods in step four. So it is likely that the email sent to the customer results in a valid reply in step seven.

First Virtual transactions uses off-line billing. First Virtual does not provide money atomicity. The actual transfer of funds in the First Virtual system is implemented off-line using traditional payment mechanism, such as with the telephone order or mail order described in Section 5.1. While First Virtual looks like a bank to the on-line consumer, First Virtual is a single merchant from the perspective of the financial infrastructure. Like a card-issuing bank, First Virtual will cancel a customer’s account after the customer refuses payment to First Virtual, but this will not make the previous transactions money-atomic.

First Virtual does not provide goods atomicity or certified delivery. The customer can receive goods and refuse payment.

Successful First Virtual transactions are isolated. Unsuccessful transactions are not. Because consumer refusals are tracked, the result of a customer refusal on one transaction depends on the outcome of other transactions. Too many refused transactions result in a refusal of service. Thus the lack of isolation is not a flaw in this system, since it is a result of a considered business strategy.

First Virtual transactions are consistent. The merchant and customer know if the customer was paid. Notice that there is not consistency in goods delivery. If the protocol were goods atomic, goods consistency would also be expected. The merchant may believe the customer has the goods and expect payment, while the customer will not pay.

After the final email, First Virtual transactions are durable. The customer cannot change her mind about the quality of merchandise after having approved a charge.

5.2.2 Security

First Virtual assumes that the Internet is inherently without security, and thus does not send credit card information itself over the Internet.

First Virtual is not secure. An attacker need only trap a packet which has the account identifier of a First Virtual account holder to receive information free. Since there are well-known locations which receive many of these packets (for example, the First Virtual Infohaus), finding such a packet is unlikely to be difficult. Thus the very lack of widespread interoperability between forms of network commerce is an advantage for First Virtual, since you cannot trade First Virtual account authorization for any other financial instrument.

Merchants can get customer First Virtual account information but not customer credit card information. Thus, merchants do get the information necessary to authorize further purchases within the First Virtual system; including charging their own customers for items they did not select. Merchants themselves will not profit from padding charges; however, they can use this information to illegitimately obtain information goods at some cost to the customer.

Merchants cannot protect the information they sell as it travels over the Internet. Attackers may steal information goods by trapping and copying information goods as they are sent to legitimate First Virtual customers.

First Virtual mitigates risk by limiting interoperability -- although the First Virtual system is not secure it isolates and limits security failures through business practices.

5.2.3 Privacy

The information available to various parties in a First Virtual transaction is shown in Table 5.2.

In First Virtual transactions the merchant gets the customer identification immediately, so merchants may easily build detailed consumer profiles. In fact, merchants are required by First Virtual to keep detailed transaction records for at least three years after the transaction (First Virtual, 1995b).

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	Partial	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf w/warrant	Full	Full	Full	Full	Full
First Virtual	Full	Full	Full	Full	Full
Observer	Full	Full	Full	Full	Full

Table 5.2: Information Available in a First Virtual Transaction

Since no messages are encrypted an observer could develop a detailed profile of consumer habits. Observers can even more easily profile a merchant's on-line business by watching only one server location.

The customer can choose a First Virtual pseudonym. If the customer takes advantage of this, a merchant can identify serial purchases by a customer but cannot link that information to any non-First Virtual transaction data.

5.2.4 Regulatory Issues

First Virtual can provide all the information necessary for any regulatory purposes. In fact, First Virtual provides more information than legally required.

Neither merchants nor First Virtual is covered under banking laws, so legal requirements that customer transactional data be maintained do not apply. The multi-year retention period for transactional data required of merchants reflects the time frame of interest to law enforcement rather than businesses, since for a given charge the customer's right to dispute is contractually limited to months.

First Virtual makes no attempt to control the use of merchant data. The open nature of First Virtual means that consumers do not have privacy from even casual observers. This makes the careful choice of and frequent changes in customer pseudonyms important.

First Virtual data retention reflects the need for broader controls on consumer records. Since the provision of consumer credit reporting is not the primary business function of either First Virtual or its merchants, their customer records are not covered under the Fair Credit Reporting Act.

5.3 Secure Sockets Layer

There are multiple versions of secure protocols for use on the World Wide Web, and many browsers. These include S-HTTP, encrypted telnet and ftp and the Secure Sockets Layer. These protocols can be used for Internet commerce, and in fact, Secure Sockets Layer is advertised as an Internet commerce tool by Netscape (Netscape, 1996). Here I am addressing the option offered by Netscape for use with its own browser: Secure Sockets Layer (Freier, Karlton and Kocher, 1996). I consider Version 3.0 here, as described in the appropriate Internet Draft.[19]

The Secure Sockets Layer is built to enable secure peer-to-peer communication over the Internet, not to enable electronic commerce per se. Electronic commerce is the most obvious, and possibly most frequently used, application of the Secure Sockets Layer. The Secure Sockets Layer is not an electronic commerce protocol, it is a handshake protocol for establishing a secure channel that can then be used for commerce. Possible uses include private and verifiable email, contract negotiation, and transmissions within or between institutions of sensitive data. The Secure Sockets Layer can be combined with other protocols which would be strengthened by the use of an encrypted channel, such as First Virtual. The use of the Secure Sockets Layer for customer address information is assumed in SET (Lewis, 1996).

The Secure Sockets Layer replaces the telephone line in the credit card transactions described above with an encrypted Internet connection.

The scope of the SSL is extremely limited: SSL offers only an encrypted tunnel through the Internet that enables the secure delivery of financial information.

---

[19]An example is Netscape's LivePayment, http://home.netscape.com/comprod/products/iapps/livepayment.html.

The Secure Sockets Layer enables traditional credit card transactions over the Internet. It is for charges sufficiently large that there is no need for aggregation. The Secure Sockets Layer requires that merchants be credit card merchants in the traditional sense: the merchant must have a merchant credit card account with an acquiring bank.

5.3.1 A Transaction

The Secure Sockets Layer protocol begins with an exchange of certificates and ends with an exchange of keys. The information transmitted in further exchanges is completely transparent to the protocol. Figure 5.3 illustrates a Secure Sockets Layer transaction. The bank is not shown because communication with the bank is off-line, i.e. not on the Internet but over a private leased line.

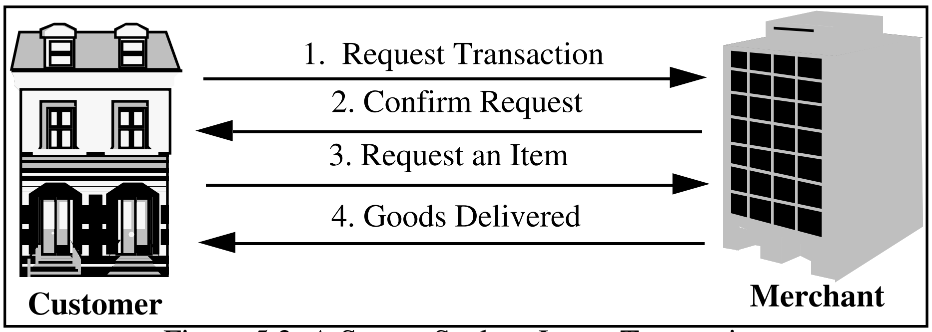

Figure 5.3: A Secure Sockets Layer Transaction

There are options within Secure Sockets Layer for authentication and key exchange for users with and without certificates. For consistency across protocols I assume the customer and merchants have certificates.

In the first two steps the customer and merchant authenticate themselves to the other and generate a shared key. There is a message for use by the customer for requesting the merchant's certificate, and thus the customer is assumed to have the certificate. The customer uses the merchant's public key to initiate a transaction. The transaction begins with key generation. The merchant replies and may request the customer's certificate. The customer and merchant use the public keys contained within their certificates for authentication and the generation of a symmetric key. Secure Sockets Layer as implemented in Netscape's Navigator assumes the customer has a certificate from Verisign, RSA or a small number of other certificate vendors (Verisign, 1996). Although authentication and key generation require more than two messages, this exchange can reasonably be modeled as two functional steps.

In the third step, using the protection provided by symmetric encryption, the customer sends her credit card number. In the fourth step the merchant should deliver the goods. These two steps are not covered in the protocol.

Notice that after step three the merchant will almost certainly obtain authorization from the customer's credit card provider, and therefore could authorize a transaction with the bank at any time after that step. Since the communication with the bank is not included in SSL this is not shown.

The SSL protocol provides for a handshake for authentication and the generation of a shared key. Thus it clearly cannot provide atomicity. Since Internet purchases are treated by credit card companies as telephone orders, the customer can refuse payment to the merchant. Thus the lack of money atomicity in the off-line financial system suggests that there is no money atomicity in a transaction using the Secure Sockets Layer. There is neither goods atomicity nor certified delivery.

Consistency, durability and isolation are as in a telephone order, as described in the first section of this chapter.

5.3.2 Security

The greatest security threat with the Secure Sockets Layer is that merchants must all keep secure servers for credit card numbers to remain secure. Thus, the customer must trust not only the merchant and his employees, but also his technical acumen in computer security. The theft of twenty thousand credit card numbers from Netcom illustrates that this is a problematic proposition. If there are dishonest employees, inadequate organizational security procedures, faulty installation of secure software, or if the merchant does not provide timely patches for his operating system, the consumer is at risk for credit card fraud.

If the merchant is honest, the employees may remain a problem. Replay attacks are trivial with the information provided to the merchant.

The effective regulatory limitation of key length to forty bits is a weakness, since the payment authorization information is not transaction-specific. Thus observers could obtain credit card authorization information, using attacks as described in Section 3.2.

5.3.3 Privacy

Because the Secure Sockets Layer provides not financial services, but rather software to create an encryption-secured connection through the Internet, Netscape has no information about the transaction. The off-line bank is the financial service provider, Netscape is the transmission security software provider, a third party may provide certificates, and the merchant initiates financial transaction authorization.

By using encryption the Secure Sockets Layer has added some degree of privacy in comparison with credit card transactions in that observers can no longer obtain payment information. (Recall Table 5.1.) This priority, hiding information from observers, reflects the fact that observation is much easier on the Internet than with the telephone network.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	Full	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf w/warrant	Full	Full	Full	Full	Full
Netscape	None	None	None	None	None
Bank	Full	Full	Full	None	Full
Observer	Full	Full	Full*	None	None

Table 5.3: Information Available in a Transaction Using Secure Sockets Layer

Identity information is available because the certificates for customer and merchant authentication are sent unencrypted.

The observer is shown as being uncertain about the date of the transaction.. This is because the observer cannot determine if a transaction actually took place -- only that there was communication between the customer and the merchant. Thus the addition of the * notation.

Information is concentrated at the off-line financial services provider, the acquirer bank. The bank has all the abilities to correlate and distribute information as in any other transaction.

5.3.4 Regulatory Issues

The Secure Sockets Layer sets up secure connections through an open network. Netscape as an entity does not have any information about what data has passed through a secure sockets connection.

In the Secure Sockets Layer the merchant keeps the payment authorization information from the customer. Thus the merchant has responsibility for protecting all customer information. This has proven problematic in the off-line payment world, with disbarred merchants and dishonest employers using credit card information to make unauthorized charges.

Finally the amount of consumer information held by the merchant and financial services providers further strengthens the argument that limits on secondary financial information need to be expanded.

5.4 Secure Electronic Transactions

With the proposed standard, the Secure Electronic Transaction protocol (SET) (Mastercard, 1996), any acquirer could have a presence on the Internet capable of authorizing and possibly capturing charges. SET is a combination of Mastercard's Secure Electronic Payment Protocol (SEPP) (Mastercard, 1995) and the Visa Secure Transaction Technology (STT)(Visa, 1995) protocol.

SET does not necessarily aggregate purchases, although a merchant may choose to ask for verification and payment in batches. This is feasible for the obvious reason that there is no need to aggregate large purchases made over the Internet. The same customer support, order processing, administration, and promotion savings that can be obtained by other purveyors of electronic commerce can be obtained by traditional credit card acquirers. The SET protocol may not compete with so much as complement the approaches of the previously mentioned Internet commerce providers.

SET models Internet commerce as mail order and telephone commerce. The merchant, rather than the acquirer, takes the risk for invalid purchases as in mail and telephone orders because a physical card is not used. However, if the secret key corresponding to an SET certificate is eventually stored in a smart card, SET transactions may be reclassified as *card present* transactions. In card present transactions the merchant is guaranteed payment by virtue of having a physical imprint of the customer's card.

Initially in SET only traditional merchants will be allowed to sell goods. This means that small publishers and professionals working at home cannot use SET if they do not have merchant accounts with an acquirer which is authorized to clear the customer's requests credit card purchases.

SET is an open standard and was developed using an open process of issuing drafts and requesting comments. Originally, Visa proposed a proprietary system with Microsoft, possibly in an attempt to leverage the dominance of the Microsoft operating system to popularize STT technology. The decision by Visa to pursue an open process is significant in terms of the promise of future interoperability. However, unlike open Internet standards Visa and Mastercard retain ownership of SET.

The weakness in terms of interoperability is that a requirement for credit card ownership sharply limits the participants in Internet commerce. In addition, the requirement that merchants have traditional merchant accounts in order to accept funds seriously limits the potential of Internet commerce by limiting the merchant population to traditional merchants. The inability of any user to charge for services may not limit the entrepreneurial power of Internet users, but it does fail to encourage individuals to create information start-ups.

The probability that SET will be a common standard is supported by simple observation of the financial strength of its founders, Visa and Mastercard. CyberCash, American Express and Europay have adopted SET as a standard.

5.4.1 A Transaction

SET offers multiple protocols for electronic commerce which reflect the different types of Internet access available. Transactions are possible for customers with email connectivity and Web connectivity. Transactions can be implemented by customers with or without certificates. Here I consider transactions when a customer has public key certificates and Web access. This is the appropriate model for maintaining consistency across comparisons of different protocols.

An alternative version assumes only that customers can calculate hash values of payment information. This protects payment information from merchants. However, this version of the protocols does not prevent replay attacks once credit card information is obtained.

In SET, the standard language in the credit card industry is used. I use slightly different language in order to be consistent with other descriptions. Credit card verification is referred to as *authorization* and payment is referred to as *capture*. Since the words authorization and capture have other, specific meanings in computer security I will continue to use the terms verification and payment. The bank as shown in Figure 5.4 is an *acquirer gateway,* a service provider for acquirer banks. That is, the gateway is the bank's Internet presence. With these changes in terminology the figure below corresponds to the description in the SET specifications.

Since SET specifications permit batching, the contents of a specific message may vary slightly from the single transaction model shown here. In fact, steps five, six, nine and ten can precede steps four, seven and eight. However, this may vary the contents of these messages in unspecified ways, so these variations are not considered.

Each description of a message that follows Figure 5.4 includes the name of the message in the SET protocol in order to simplify further exploration of the SET Technical Specifications for the reader. In addition, corresponding names for variable descriptions are given in parentheses. This is not done in other protocol analyses because the explanation of the other protocols is available in a simpler format than the SET Technical Specifications.

A transaction using the SET protocol for an interactive medium is show in Figure 5.4. Notice that browsing and price negotiation are not included in the SET protocol. A corresponding diagram can be found on page 133 of the SET Technical Specifications.

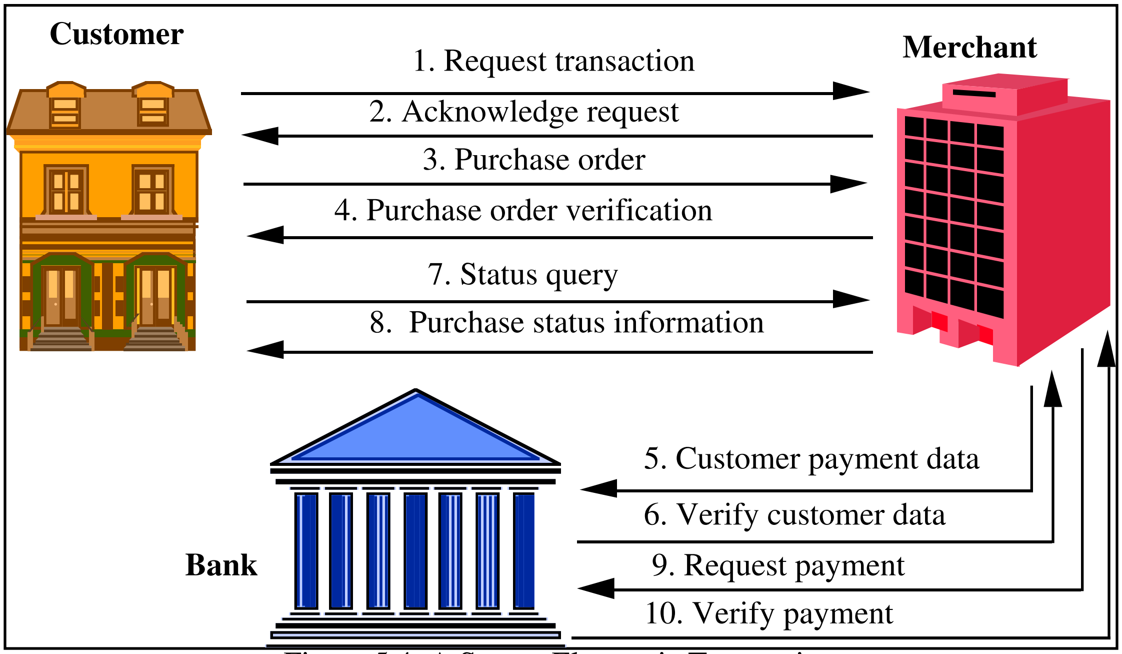

Figure 5.4: A Secure Electronic Transaction

In the first message (PInitReq) the customer identifies her desire to make a purchase. This message includes a customer-specific message identifier (LID_C), a corresponding nonce (Chall_C), the customer selected payment method (BrandID) and a list of certificates with the appropriate hashes for verification. There is no encryption used to protect this information; however, one presumes that only the cardholder has any interest in sending the cardholder's certificate. Thus after this step the merchant, and an observer, know the customer's identity, the merchant to whom it was directed, the item requested and the price.

In the second message (PInitRes) the merchant acknowledges the customer's request to begin a transaction. The merchant begins a record which includes the customer's transaction identifier and brand in the database. Presumably the customer's address (for responses) is also included, although this is not noted in the specifications. According to the documentation, the customer address is supposed to be obtained *out of band* by the merchant. Thus the message which contains this information is not specified, although the Secure Sockets Layer is an obvious choice. The merchant will now know the customer's credit card type, limits on the customer's account, and any customer attributes implied by this credit information. The merchant must know the customer's billing address for authorization. Thus, although the method for obtaining that information is not specified, the data are reasonably included as part of the information exchange in SET.

The second message includes the transaction identifier to be used by all three parties, a corresponding challenge, a time stamp, and a new challenge from merchant to customer. After this message the customer has the merchant's certificate.

The third message is the customer's purchase request. This is the customer's conditional commitment to completing the transaction. Note that the customer is not committed, and the payment is not durable, for some weeks. The customer maintains the right of refusal until after she reviews her monthly credit card charge account summary.

The purchase request is the most complex message: it includes payment and order information. The payment information is encrypted so that the merchant cannot read it but the bank can. The hash of the order information is included in the signed message. The order information and purchase amount themselves are sent in verifiable but unreadable form, i.e. they are hashed. The order information is obtained by the merchant external to the protocol. This message is signed by the customer. The message includes a general description of goods, amount, and nonces in the clear. The payment information includes account number, transaction identifier, amount and card expiration date encrypted for the bank.

The fourth message is from the merchant to the customer. Note that the merchant may choose to obtain authorization, or to respond to the customer immediately and batch authorization. Here I have assumed the former. If it were the former, the merchant would send the results of his attempt to obtain authorization.

The fourth message is the merchant's indication to the customer that the merchant will complete the transaction, contingent on authorization, and possibly capture. This message is signed by the merchant. It includes the transaction identifier, the customer's transaction identifier, and a flag indicating the status of the transaction.

If authorization had been completed previously, then the authorization status and authorization amount would be included. If the transfer of funds, called capture, had been completed, the capture status, capture amount, and the ratio of amount captured to purchase price would be included.

The fifth message is the authorization request from the merchant to the bank. (Recall that the bank here is actually a gateway to the credit clearance system.) This message is encrypted so that the bank can read it. The message is signed for verification, then encrypted using a one-time DES key. The DES key itself is then encrypted in the bank's public key. The same technique is used for the capture message to and from the bank; although the authorization response does use 'extra-strong' encryption. The definition of extra-strong is not specified in the SET documentation, but observation of current bank practice and later communications allow me to assert that 'extra-strong' means triple-DES.

The authorization request includes transaction-specific and merchant-specific data. Transaction-specific data includes the transaction identifier, the date, and the order information. The order description, the amount, and nonce are hashed together and included. The transaction identifier, the customer's transaction identifier, the date, merchant identifying information and the brand identifier are hashed together for inclusion.

The merchant-generated data include the amount, the merchant's business area, and one byte identifying a specific purchase area. This single letter is referred to as the MarketSpecData and identifies the industry -- hotel, auto, etc.

The customers' address is included. There is also an option for requesting additional authorization, above the purchase amount, called AdditionalAmount. Finally there is a flag to identify the message as part of a batch, and fields for associated batch information. The merchant includes the customer's billing addressed in an authorization request.

The sixth message is the authorization response of the bank to the merchant. Before responding the bank verifies the signature and verifies signatures of the hash values signed by the customer and sent by the merchant match. The authorization response is encrypted with "extra-strong" encryption. This message is signed by the bank.

In step seven the customer may contact the merchant to determine the status of the transaction. Only the transaction identifier and the customer's transaction identifier are in this message.

In step eight the merchant responds with a signed message of the same form as message four. That is, the merchant reiterates his commitment to complete the transaction and notifies the customer of the authorization status.

In step nine the merchant requests capture from the bank. Capture follows authorization. In authorization a certain amount is reserved on the credit line of the customer. In capture, a lesser or equal amount is transferred to the merchant. Capture is reversible in Internet, telephone and mail order transactions from the perspective of the merchant. The capture messages are both signed and encrypted using a one-time DES key, which is then protected using the recipient's public key.

In step nine the merchant sends the transaction identifier, the date, transaction-specific data (from the authorization request), and amount. There is also additional data for ease of processing if the order is batched. This message is the merchant's commitment to the bank to complete the transaction.

In step ten the bank confirms capture.

Consider the transactional characteristics of SET.

The SET protocol does not assure isolation because of the inclusion of an AdditionalAmount field in the authorization request from the merchant to the bank. The customer neither approves nor even has knowledge of this field. This has proven problematic with physical card transactions only occasionally as described in Section 5.1.

The lack of isolation may be exacerbated by the ability of electronic customers to travel between merchants at a much higher rate than physical customers. If a consumer visits many Web pages and each blocks off an amount that assures maximum possible payment, the consumer may quickly be drained of available credit.

Credit card transactions are normally consistent in that the customer and merchant agree on the amount paid.

Credit card transactions are durable, after the customer has agreed to a charge.

The SET protocol provides money atomicity, but does not provide either goods atomicity or certified delivery. The SET protocol could be strengthened by the addition of certified delivery for information goods.

In the following paragraphs I offer one technique for providing certified delivery; certified delivery could be added in other ways. This could be done by adding a message where the information is delivered encrypted, and then adding the appropriate fields. The following paragraphs describe the changes necessary to provide certified delivery for SET purchases. The additional messages, three and four, are shown in Figure 5.5.

Of course, there are several possible ways to add certified delivery to SET. For example, one could add fields to the initial two steps so that the merchant responds immediately to the customer's request. However, the fields in the SET protocol are clearly specified and there is an acceptance in the protocol for information to be received external to the protocol.

Thus the addition of a new message and the alteration of the fields appears to be a preferable choice.

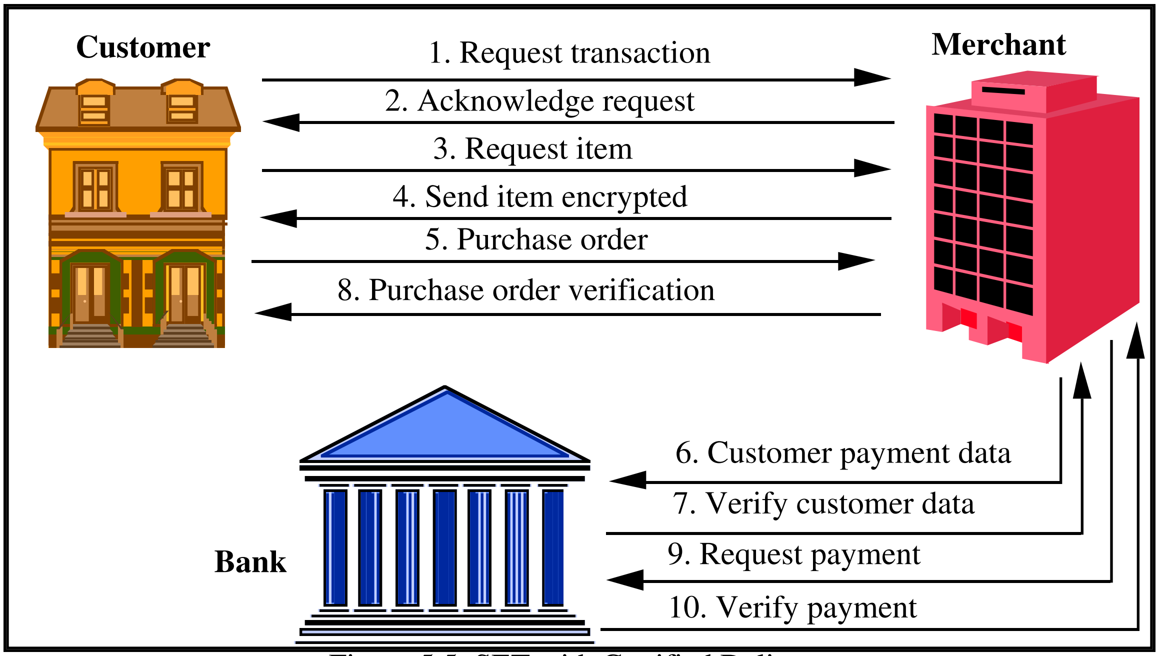

Figure 5.5: SET with Certified Delivery

The significant changes are that the purchase order verification (Pres) must follow the authorization requests. This means that certified delivery purchases cannot be batched. This is a variation already defined within the SET protocol, so this is not a change to the standard.

Between the order request and order information the merchant must determine the item the customer wants, encrypt it, and send it. A version of the inquiry request can be used to request an item from a catalog; however, more complex order descriptions can provide greater protection for the customer.

The delivery of the encrypted merchandise is an additional step required for certified delivery. This may be transparent or external to the SET protocol, since it is the customer's acknowledgment of the item received that is critical.

After the second message the transaction ID has been established. The customer needs to send a message with the item requested. This message must include the item ordered and the price.

The merchant would send the item encrypted with some key, k. This message must be signed and encrypted. The message must be signed for purposes of nonrepudiation. If the message is not encrypted, then the customer could observe the transactions and obtain the key. The message would be as follows, in the SET format: S_M{TransID, H(E_k(item))}, E_k(item). Thus, the merchant would have signed a commitment to delivering an item with a specific description, acknowledgment of having delivered a specific bitstream, and that this encrypted bitstream is the item described, encrypted with the enclosed key.

The customer will have the encrypted merchandise when building the purchase order. (Step five in Figure 5.5; Step three in Figure 5.4.) The customer already has the merchant's signature on a description of the item and the price, as described under Figure 5.4. The customer would also include a hash of the encrypted item sent by the merchant. The order description field (OD) can be expanded to include this hash value. After this step, in addition to the information discussed above, the merchant has nonrepudiation from the customer that the bitstream was received.

The merchant may then ask for authorization. This message must be altered to identify the transaction as information-based and include the key. The next paragraphs describe the field used to enable this. The purposes of the data in this message are to provide the bank with the key to the merchandise, and the customer's verification that the merchandise has been received. Of course, the payment authorization information as defined in SET must remain.

The merchant's verification of the customer's payment information must also be changed. (This is Step five in Figure 5.4 and Step six in Figure 5.5.) The changes are in the data fields called the Authorization Request Payload. First, the transaction must be identified as one concerning an information good. The natural place for such a specification is the MarketSpec field in the MarketSpecData (Mastercard, 1996, p. 130). Recall that this field identifies the industry. For certified delivery the MarketSpecData field could be *I* for information.

The SaleDetail is also part of the Authorization Request Payload. SaleDetail includes information on specific transactions. The fifth field of SaleDetail is either PassengerTransport, LodgingDesc or CarRentalDesc. An option for certified delivery would be to include the key in this field. If the MarketSpecData was *I*, then the merchant would be required to include the key to decrypt the message for certified-delivery messages.

The information is structured as follows in SET:
AuthRequestPaylod(MarketSpecData[MarketSpec], SaleReqData [{SaleDetail(Desc)}])

All of this is contained in the AuthReqData, which also contains the hash of the order description, the purchase amount, and a nonce. As such:
AuthReqData == {TransID, AuthReqData, H(OID), H(OD, Amount, ODSalt), PI, AuthReqPayload, Thumbs}

The Authorization Request Payload is contained in Authorization Request Data. The other field of interest in the Authorization Request Data is the order description (OD). The order description contains the hash of the item, as previously described. The order description is signed by both the merchant and the customer.

The Purchase Order Verification can include the key. (Step ten in Figure 5.4; Step eight in Figure 5.5.) The CompletionCode can include the key if the field is sufficiently large. This cannot be determined with current documentation.

Thus leaves one question remaining: how does the customer obtain the key if the merchant fails? Having the customer obtain the key from the acquirer gateway would change the functionality required of the gateway system. I consider two other options: having a dedicated key server or providing the information through the certificate authorizing agent.

Having a dedicated key server has the advantage that the server could be widely available. It would have the disadvantage that it requires creating an entirely new entity. The dedicated key server can be administered by the gateway system, the acquiring bank, or by some third party trusted by the merchant and the customer. The keys could be made publicly readable, since the key itself, absent the merchandise, is not valuable. There is no additional threat of theft by making keys readily available.

Using the certificate server has the advantage that customers already interact with the server in question, and with the addition of a key query the certificate server can also be a key server. The certificate server has two characteristics that make it a promising alternative: it is readily available to the customer and merchant base, and it is closely connected to the acquirer gateway to verify certificate information. Since the certificate-granting authority must be tightly bound to the financial system to avoid presentation of faulty credentials to obtain a certificate, connecting it to the gateway (or on-line presence of the bank) does not create an additional security risk.

Suppose that the certificate authority can service key queries. This is a much simpler message than a certificate request, or certificate renewal request. Although there is more than adequate space in the certificate request messages, a smaller key-specific query would be a reasonable addition.

The merchant must have selected a certificate server in order to complete a transaction. The customer can prove that the merchant selected a particular certificate server by presenting the merchant's credentials. Thus a customer can prove that a merchant did not provide a key as promised by presenting the merchant's certificate, the contract, the encrypted merchandise and a response from the appropriate certificate authority indicating that there is no corresponding key available.

Notice that this extension does not require the creation of new trusted entities. The merchant is vulnerable to fraud because the certificate authority could claim not to have the key, yet transmit the key to the customer. However, the perversion of a certificate authority would result in the ability to create apparently valid merchant and customer certificates, which has much more severe consequences.

5.4.2 Security

The most dramatic improvement of the Internet protocol over the mail order and telephone protocol for Mastercard is that the merchant gets enough information for only one purchase. Merchants cannot use SET information for replay attacks. Not only are transaction identifiers not repeated, the date and time of the transaction are included. A merchant cannot produce a purchase order signed with the customer's private key with different time and transaction identifiers.

The SET protocol does not include negotiation or verification of delivery of information goods. A customer can claim not to have received goods already consumed, and a merchant could claim to have provided goods not sent. Therefore the security of SET depends upon the delivery mechanism used. The strength of nonrepudiation is limited when the promise can be verified, but the fulfillment of the promise cannot be.

The lack of goods atomicity creates the potential for fraud. The addition of certified delivery can address this for information goods; however, there is no such solution for physical merchandise.

The protocol includes the possibility for a pseudonym in terms of the account number. That is, the customer can choose to use a fake account number. Since the possession of an account number creates the possibility for fraud, there is no reason that this should be an option. Requiring that all account numbers are pseudonyms is a low-cost technique for increasing security. By having pseudonymous identities linked to the pseudonymous account numbers, the interest of privacy would be served as well.

Customer address and order information are provided in a separate channel from SET. Thus this information is available to observers. How problematic this information leak is depends upon the importance of customer base information to the merchant and the importance of transactional information to the customer.

Customer address is information used for verification. Thus, one element of verification information is sent in a way that is neither secure nor private. This is similar to the separate channels used for purchase and verification in First Virtual.

5.4.3 Privacy

The information available in a SET transaction is shown below in Table 5.4.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	Partial	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf w/warrant	Full	Full	Full	Full	Full
Bank	Full	Full	Full	Full	Full
Electronic Observer	Full	Full	Full	Full	Full

Table 5.4: Information Available In a SET Transaction

The SET protocol provides more privacy than current credit card transactions, since the customer can choose a pseudonymous account number. This implies that the capacity for using pseudonyms is built into SET, although it is not currently explicit. Note that the fact that financial information is hidden from the merchant increases security, not privacy.

The electronic observer has complete knowledge about the transaction because the certificates containing identity information are transmitted in the clear. Encryption is used to obscure payment information, not order information. SET messages could be sent over an SSL connection to protect information from observers.

Recall that the merchant knows not only the customer identity but also customer attributes, including address. This information has the potential to be more than just a privacy violation. The availability of this information, and the ability to correlate it in real time with other ethnographic and economic data create the potential for electronic red-lining.
SET is less private than NetBill since the bank (through the acquirer gateway) knows the item(s) purchased. It is more private than First Virtual, since the merchant is not apparently required to maintain records of customer purchase for three years. This is made unnecessary by the nonrepudiation enabled with public key cryptography, the contractual limits on merchant loss, and the statutory limits on customer loss.

5.4.4 Regulatory Issues

SET is an open system that provides all information necessary for regulatory purposes. SET offers very little privacy, primarily because of the inclusion of the customer’s name and address as information required from the merchant. With the use of certificates and public keys the security advantage of including such information is questionable for items not requiring physical delivery. In fact, the use of a pseudonymous certificate with no physical customer information would require no change in the protocols and would offer a vast improvement in consumer privacy.

The concentration of information at the bank, actually an Acquirer Gateway, in this system reinforces the need to extend the constraints on secondary use of information beyond credit reporting agencies.

The stated reason for limiting privacy according to the SET Business Strategy is the desire to export this protocol and the current controls of the export of cryptography. This regulatory-driven limit on privacy reflects a need to recognize that transactional information correlated with identity is itself valuable and worth protecting. The limit on cryptography to the purely financial weakens the security of the SET system, since information which itself is not strictly financial but which would be useful for identity theft is sent in the clear.

5.5 NetBill

The Carnegie Mellon Internet billing server (Sirbu and Tygar, 1995), NetBill, http://www.netbill.com/, uses an electronic ledger system, a bank which holds all money, and customers and merchants who send authorizations for transactions. NetBill aggregates transactions, resolves disputes, and sends account transfer instructions to the bank.

Like First Virtual, NetBill is optimized for the purchase of information goods on-line. NetBill would serve the same market as First Virtual: low value, quickly consumed information goods. NetBill is designed to reduce transactions costs by using the factors which make network goods difficult to purchase. NetBill is designed so that any consumer with a bank account can be an information provider using NetBill merchant software. Thus NetBill has the potential for serving all possible information merchants on the Internet.

NetBill is designed to be able to profitably sell low price goods. NetBill also has a simplified protocol for free goods. Examples of zero priced goods are additional issues after the purchase of a subscription, or targeted coupons. The simplified protocol enables merchants to distribute these goods without being concerned that they will be made available to observers during transmission.

NetBill provides aggregation services as an intermediary. Aggregating transactions of ten and twenty cents into ten and twenty dollar charges results in orders of magnitude of cost spreading. NetBill automates authorization, customer service and much dispute resolution. Promotion is also automated, since NetBill is aimed at the on-line consumer. Unlike First Virtual, NetBill can also make high value sales. NetBill provides security and business records necessary to provide the dispute resolution required for high value purchases.

NetBill reduces the cost of account acquisition and credit processing by accepting standard methods of payments through banks. NetBill provides per-transaction authorization and transaction aggregation using customer credit card or bank accounts.

The business plan of NetBill also makes clear that the provision of clients, servers and transaction processing should be subject to competition. By using open standards, NetBill can prevent any one server or software provider from becoming a bottleneck. Similar considerations drove Mastercard to create a system based on open standards (Mastercard, 1995) and Visa to join in that effort (Mastercard, 1996).

Because NetBill can provide verified orders, this protocol could be used to provide verifiable receipts for orders of physical goods over the Internet. This would require the use of current verified delivery techniques, such as registered mail, for physical delivery. Currently, the extension of NetBill for verification of purchase orders for physical goods is not under consideration. Clearly NetBill cannot provide certified delivery for physical goods, as there is no economically feasible way to verify the exact physical good delivered.

5.5.1 A Transaction

NetBill transactions require eight steps. In the description of the steps below recall that each party, x, has a public and a secret key, denoted as X and x respectively. Shared symmetric keys for the session are denoted as xy, where x and y are the parties that share the key. For example, the shared key generated in the customer's Kerberos ticket is denoted cm since it is shared by the customer and the merchant. The Kerberos ticket, which contains verification of the sender's identity, is denoted as "sender ticket". Table 5.5 shows the definitions of terms introduced in the protocol description.

1: C -> M customer ticket, E_{cm}(credentials, product order, request flags, QID)
2: M -> C E_{cm}(product id, price, item description, request flags, QID)
3: C -> M customer ticket, E_{cm} (QID)
4: M -> C E_k(goods), E_{cm}(TID)
5: C -> M customer ticket, E_{cm}((electronic purchase order $)_c$)
6: M -> N merchant ticket, E_{mn}((electronic purchase order$)_c$, merchant account, merchant memo field, k$)_m$)
7: N -> M E_{mn}((receipt$)_n$), E_{cn}(TID, customer account, account balance, flags)
8: M - > C E_{cm}(receipt$)_n$), E_{cn}(TID, customer account, account balance, flags)

Field	**Description**
product order	Machine readable description of the goods being ordered.
item description	A human readable description of the goods being ordered.
QID	Quotation identifier. Unique within merchant. Allows merchant to associate positive response to quote with information sent in request for quotation.
TID	Transaction identifier. Globally unique identifier. Includes a timestamp.
request flags	Flags specifying various delivery options.
electronic purchase order	Customer's identity, item description, the merchants identity, the price, the hash of the encrypted goods, the hash of the product order, and the TID. Also customer account information and customer memo field encrypted with shared key cn (i.e. unreadable by the merchant).
customer memo field	Text field entered by customer--similar to memo field on a check.
merchant memo field	Text field entered by merchant.
receipt	Customer identity, merchant identity, price, TID, k.
flags	Customer account status information.

Table 5.5: Fields in the NetBill Protocol Definition

The steps in a NetBill transaction are shown in Figure 5.6.

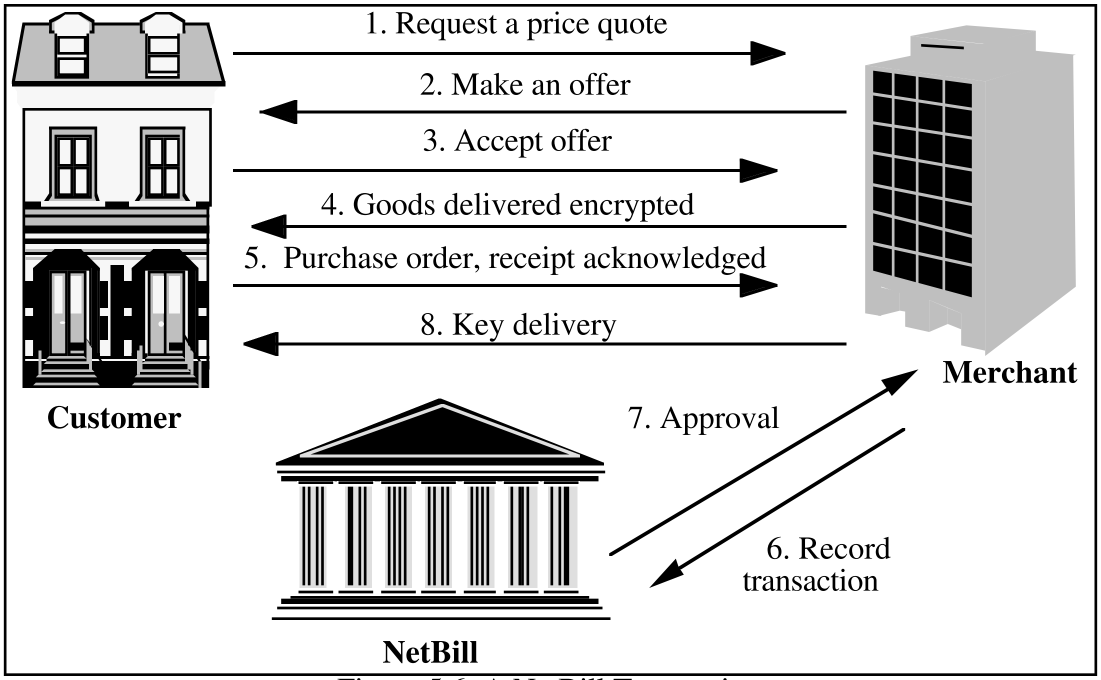

Figure 5.6: A NetBill Transaction

Any time after step two the customer still has the option of buying the selected item at the offered price. After step three the merchant has a promise to buy, but not a promise with nonrepudiation. Until step five the customer can abort the transaction. If this protocol is interrupted at any time before step five it can be restarted without loss.

Note the merchant does not encrypt the goods with the shared key, cm. The merchant can repeatedly send the item without fear of theft because the key to the goods is sent encrypted, in the receipt.

Notice that the customer cannot verify that the item delivered is indeed the item requested until step eight. At step four, some encrypted item is delivered, but the customer cannot know if it is the item promised. By endorsing the electronic purchase order, containing a checksum of the goods, the merchant is acknowledging that the goods received by the consumer are what the merchant intended to send. NetBill receives the endorsed checksums of the purchase order and of the item delivered.

After step six NetBill knows not only that some transaction has occurred, but also that the two account changes required in the transaction are linked. This means that NetBill transactions are money atomic.
NetBill is goods-atomic, and provides certified delivery. NetBill is goods-atomic because if the customer does not receive the item, the merchant cannot obtain a signed purchase order. NetBill is goods atomic because the information contained in the purchase order is signed by the customer after delivery, thus the customer cannot falsely deny receipt and the merchant cannot falsely claim delivery.

NetBill provides certified delivery because the description of merchandise, the key, and the item delivered can all be verified from step six. The purchase order includes a description of the item requested and a checksum of the item delivered, both digitally signed. If upon decryption, the item delivered does not match the description, then the customer will get a refund. Note also that the merchant must sign and register the key to obtain payment. Thus the customer is assured key availability upon payment.

In addition to account changes being atomic, they are consistent. The merchant and customer agree on the amount, the item delivered, and the item promised.

There will undoubtedly be quality disputes. Recall that First Virtual solves this problem by automatically refunding money and monitoring consumers. Conversely, NetBill monitors merchant complaints. For example, suppose a customer wants the papers describing the NetBill protocol, and orders “NetBill papers.” The customer may receive the policy descriptions and business plans of NetBill, but without the specific encryption schemes and communications protocols. The description may have been correct, “the NetBill papers,” but the information the customer presumed was there, and specifically wanted, was not included. Thus certified delivery is a powerful, but semantic claim. You get what you negotiate.

5.5.2 Security

NetBill depends upon the security of customer and merchant keys, as well as the security of the NetBill servers.

If a NetBill customer loses her key the attacker could purchase information and charge the customer. The attacker could spend all the money currently in the customer’s account. However, the attacker could not transfer any additional fund into the customer’s account, since the transfer of funds requires a challenge and response sequence in addition to cryptographic keys.

If the attacker had access to merchant keys and customer keys, the attacker could promise deliveries at any price. Limits on the amount of purchase will constrain the amount a customer can lose but would not necessarily help the merchant.

If a merchant key is subverted the attacker can also attempt to make fraudulent purchases but will be unable to forge a customer’s signature. Obtaining funds from a merchant account would require approval from both the merchant and the merchant’s acquirer bank. Thus an attacker is prevented from draining the merchant’s account by the business policies established between NetBill and the banking infrastructure.

If a NetBill server is subverted, the NetBill attacker could have up to one month to change accounts and abscond with funds. This is because it could take one account-activity reporting cycle for the first customer to complain of unauthorized debits. If merchants are not credited until customers approve transactions, then loss of server security would be without cost. The financial security of the NetBill server depends on funds availability policies which are yet undetermined. In this case, NetBill keeps sufficiently detailed information for recovery from a fraud so that liability could be reliably assigned.

If the attacker subverts NetBill and also obtains a customer account the attacker could obtain free merchandise. The attacker would merely reject message six, refuse payment, and keep the key to decrypt the merchandise.

5.5.3 Privacy

The information available in a NetBill transaction is shown in Table 5.6.

NetBill customers have very little privacy from NetBill, but can purchase their privacy from merchants. A customer can choose to purchase the service of a pseudonym provider. With the use of credentials, consumers may remain pseudonymous and still obtain any earned discounts. Regardless, the only information not available to the NetBill server is the item(s) purchased. NetBill knows the parties, date and amount of all transactions. Neither NetBill merchants nor NetBill servers are prohibited from compiling and selling customer information.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	Partial	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf w/warrant	Full	Full	Full	Full	Full
NetBill	Full	Full	Full	Full	None
Observer	Full	None	Full	None	None

Table 5.6: Information Available In a NetBill Transaction

NetBill is a medium privacy system. NetBill provides pseudonyms for consumers that can be used for a single transaction, or for each transaction with a particular merchant. Pseudonyms can be linked with authorization to specific discounts, and access control (for children, for example) can be maintained.

Not e that for any customer to have ready access to merchant's certificate any observer must also have access to merchant's certificates and the information contained within. Since the customer uses the merchant's key to send her own certificate, the observer cannot obtain the customer's information.

NetBill has proposed that the problem of customer location information being provided to merchants can be addressed by the use of intermediaries (Cox, 1994). Since public key transactions are used, the intermediary will have information about the content of the transaction. By using different intermediaries, the customer can both distribute billing charges and protect her identity from the merchant. Customers may also use a pseudonym; however, continued use of the same pseudonym can result in the customer's identity being part of the information connected to the pseudonym. (Chapter 2 has further discussion of pseudonyms.)

Note that even though NetBill can prove an item was delivered as promised, NetBill does not know the item in question. However, because NetBill does not include in the hash value a salt of the items, NetBill can implement a dictionary attack, where all possible values are hashed and then the customer's order compared to these known values. This is enabled by the use of checksums, or secure hash algorithms, as described in Chapter 2.

5.5.4 Regulatory Issues

NetBill provides all information necessary for regulatory purposes. If NetBill required that merchants keep information about consumer purchases, then there would be complete information for law enforcement with the cooperation of NetBill and the merchants.

Despite the lack of exact information on purchases, NetBill does keep fairly detailed information of customer purchasing habits. There is no constraint on NetBill selling this information without the customer's knowledge or consent. This reinforces the policy conclusions that controls on the distribution of consumer financial data are too narrowly focused.

5.6 Anonymous Credit Cards

Anonymous Credit Cards are, as the name implies, based on a credit model. The Anonymous Credit Card protocol would appear to have high transactions costs, since so many parties and messages are required. Thus the Anonymous Credit Card protocol assumes that customer will pay the cost of higher transactions for conditional privacy. The Anonymous Credit Card protocol also assumes that law enforcement will pay the price of difficult data recovery.

To protect privacy the protocol provides conditionally pseudonymous bank accounts. To provide anonymous communication the protocol provides the service of a centralized exchange, CX. This is the only protocol which explicitly addresses the loss of privacy in the communications process in every purchase.

5.6.1 Transaction

In the Anonymous Credit Card protocol each customer has two banks: an account at the customer's home bank, B_C, and a pseudonymous account at a second bank, B_P. The pseudonymous bank, B_P, believes that the customer's home bank, B_C is creditworthy. The customer's home bank believes that the customer is creditworthy and thus extends credit to the pseudonymous bank upon the customer's request. The merchant has an account at his acquiring bank, B_M. The customer's home bank knows the customer's identity and provides billing settlement. The pseudonymous bank is the consumer's electronic psuedonymity provider.

A transaction begins when a customer selects an item. The customer would send a series of messages through the communications exchange to complete a purchase. If the merchant can match the browsing to the transaction, identity information is leaked as previously discussed. Negotiation is not included in this protocol.

The information distribution in the transaction is described below. Note that messages seven, eight and nine also go through the communications exchange. In each of those message exchanges the only information obtained by the communications exchange is that someone at the customer's address is communicating with the bank.

The first message is not specified and need not be on-line, since it is infrequent. Presumably the customer could use any number of currently available methods for transmitting instructions to a bank.

After the second message the only information that has been exchanged is that the communications exchange knows that the customer's bank wants to contact the pseudonymous bank. Note that the customer's bank does not know which bank is the pseudonymous bank, since that information is encrypted so that only the communications exchange can read it. The communications exchange cannot read the contents of the message being transmitted from the customer bank to the pseudonymous bank because it is encrypted with the pseudonymous bank's public key.

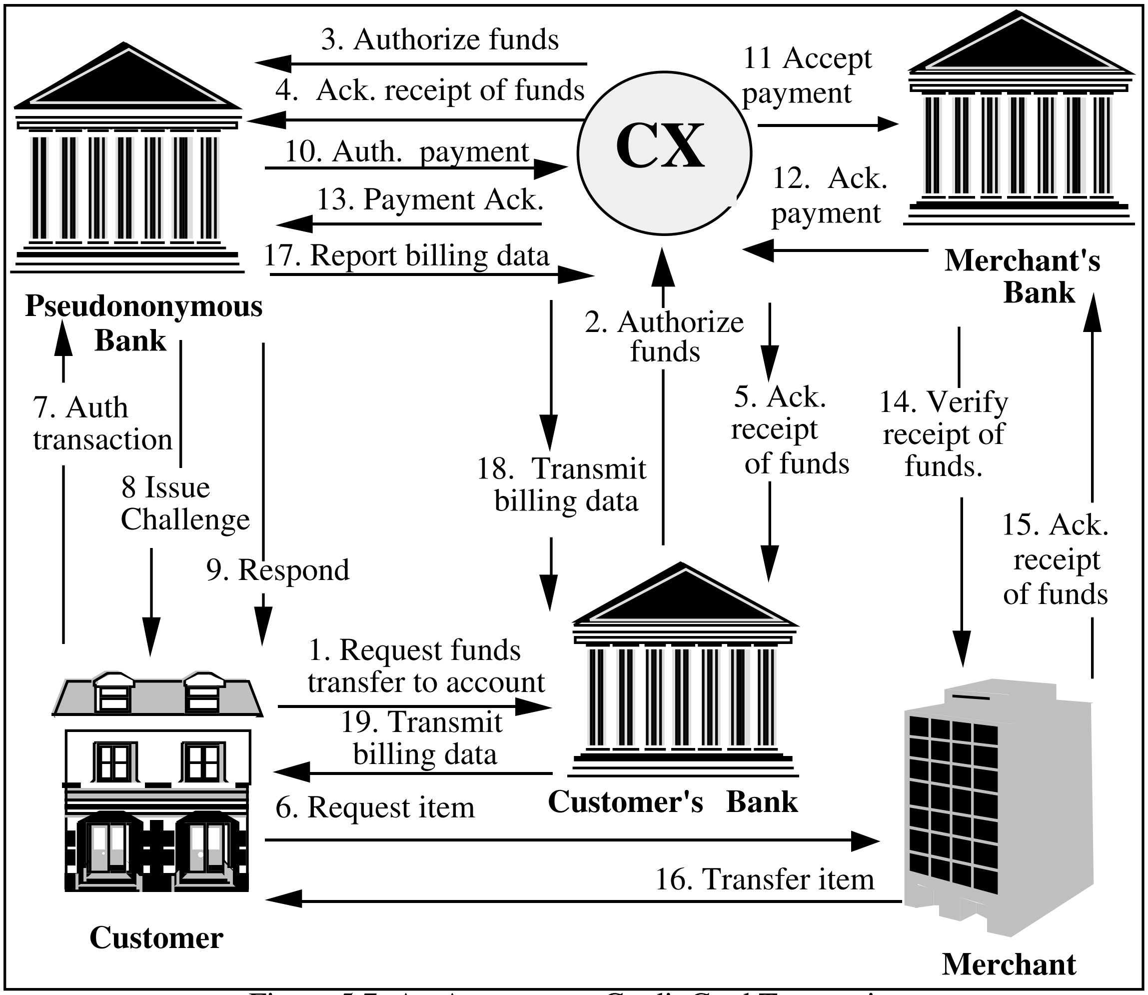

Figure 5.7: An Anonymous Credit Card Transaction

In the third message the communications exchange forwards the message to the pseudonymous bank. The pseudonymous bank verifies the request and extends credit to the customer under the customer's pseudonym.

After the fourth message the communications exchange knows that the customer's bank successfully contacted the pseudonymous bank.

In the fifth message the customer bank receives verification that the extension of credit has been completed.

In the sixth message the merchant receives a request for an item. At this time the customer and merchant agree on the item and price. As for message six, there is actually no discussion of this exchange. Presumably the customer obtains the merchant’s public key certificate to get the merchant’s public key. Also the merchant must send information on the merchant’s bank, including his bank account. Also presumably the merchant obtains the information necessary from the customer to assure delivery.

In step seven the customer sends a signed request for a transfer to the merchant’s bank account to the pseudonymous bank. This includes a personal identification number for pseudonymous authentication. Messages seven and nine are signed by the private key of the customer asymmetric key, so presumably the key in question is actually a pseudonymous key.

In message eight the bank responds with a challenge. Since this is a pseudonymous system authentication cannot rely on verification of identity. Thus for further assurance that the sender of message seven is authorized to request funds from the pseudonymous account the merchant asks a question or series of questions. These questions are determined when the account is established. The customer selects a list of questions, and also provides verifiable hash values of the answers.

In message nine the customer provides the answer to the question(s) transmitted in step eight. At this point the bank will commit funds on the customer's behalf.

In message ten the communications exchange receives the message from the bank. In this message the bank commits to providing funds to the merchant's bank. In the eleventh message the communications exchange forwards that message to the merchant's bank.

In message eleven the merchant' s bank receives the commitment of the pseudonymous bank.

In message eleven the merchant's bank receives the customer's commitment to the purchase, as communicated pseudonymously by her bank.

In the next three messages, twelve, thirteen and fourteen, the merchant bank confirms its receipt of funds. The merchant acknowledges receipt of the customer's commitment and transmits that commitment to the merchant.

In message fifteen, the merchant commits to the transaction by acknowledging receipt of the customer's funds. At this point the merchant can prove that the merchant has accepted payment.

In message sixteen, the merchant delivers the item. Note that this message is not explicit in the protocol, but is assumed. This message is put as message sixteen because this is the last transaction-specific message.

Messages seventeen through nineteen are periodic messages aggregated over multiple transactions.

Messages seventeen and eighteen consist of the delivery of the billing statement. The pseudonymous bank creates periodic billing statements, each of which is forwarded through the communications exchange to the customer's bank.

The customer’s bank then forwards the billing statement to the customer. The customer then has the opportunity to contest particular items.

The financial intermediaries and the communications exchange can share information for dispute resolution. The entire set of information available includes elements necessary for dispute resolution for money atomicity, depending on the timing of funds availability to the merchants. There is neither money atomicity nor certified delivery.

The existence of money atomicity and failure to provide money atomicity can be illustrated by considering a failure at each step. If steps one through three fail, then no funds have been exchanged.

Notice that steps one through five are themselves a transaction: the customer is purchasing pseudonymous spending authority. Thus these steps by themselves should be atomic. If the pseudonymous bank did not receive an authentication from the communications exchange by that time, the deposit would be canceled and the funds released at the customer's bank. This transaction needs a point of serialization, with global commit and the ability of all parties to inquire as to the status of a transaction.

If steps four or five fail then money would be created at the customer's pseudonymous account but not decremented at the customer's identity-linked account. This discrepancy would not be discovered until the end of the billing period. Then it would be discovered when the customer refused to pay. This requires that the customer disclose her identity -- an example of the trade-off between atomicity and privacy.

This protocol is not atomic, although there is recovery failure built in. The delay in the recovery factor creates opportunities for fraud. This system could be made money atomic if the store did not get funds availability until the customer had approved the bill. This would move the cost of fraud away from the customer to merchants, as is the case in credit card and checking transactions.

The customer bank will not detect the failure until it is time to balance accounts. Then it may be problematic determining which transaction actually failed. Neither the customer nor the merchant has an interest in identifying the misappropriation, because neither will gain when consistency is restored. This could be the most expensive failure to detect and could require costly data matching.

Now consider the purchase. If message six fails then there is no transaction. Notice for issues of privacy that the distribution of customer information through browsing occurs in this unspecified step.

If steps six through eleven fail then there is no transaction.

If either step twelve or thirteen fails then the merchant bank can verify that a transaction has occurred, while the customer bank can verify that the same transaction did not. At reconciliation this would be detected.

If steps fourteen or fifteen fail, then the pseudonymous and merchant banks believe the transaction has occurred while the merchant and the customer do not. This will be discovered during the billing process. This could be changed by requiring that steps fourteen and fifteen precede step twelve.

One general failure illustrated by these specific failures is that there is no simple way for the customer to poll the pseudonymous bank to get information about the transaction status. Any one of these message failures results in temporary creation or destruction of funds, and requires a complex, multi-party and therefore expensive resolution process.

If step sixteen fails then the customer may object on her bill. Of course, the same is true if sixteen does not fail, so this system is neither goods-atomic nor does it provide certified delivery. There is no way for the merchant to verify that goods were sent to the customer. The customer has the ability to contest payments and can prove that she paid for the goods with the assistance of the banks and the communication exchange.

The merchant has some transaction identifier which is linked to customer payment. The merchant cannot prove that the item delivered was the item requested. Only the customer has access to verifiable information about the order. If the customer began the transaction with the intent to deceive then fraud would appear straightforward. Especially on information goods, or goods the consumer may want only for a short time, this could be tempting. Conversely, there is no way to prove that the merchant did not send the goods.

If steps seventeen through nineteen fail the customer is left without information about her account.

The system can assure money atomicity, consistency, isolation and durability in the long term if the business policies are appropriate. However, the conflict resolution process is costly and requires a loss of customer anonymity.

5.6.2 Security

The security assumptions in the Anonymous Credit Card protocols are as follows: the secret keys of each party are hidden and traffic analysis is not possible. If an attacker could obtain the secret keys of any party the attacker could pose as that party. If the attacker obtained the secret keys of a bank, then that attacker could verify multiple customer accounts and siphon off their funds.

However, any attack would be detected when it came time for reconciliation of accounts.

The customer's key is undoubtedly going to be the weakest link in any security chain. In this case the loss of the customer key would appear to cause no harm since there is a challenge and response requirement. Suppose the customer must identify herself by something more than an asymmetric key when authorizing transfers to the pseudonymous account. The existence of the pseudonymous account would require an attacker to set up an additional account and then drain the funds off in a series of transactions. If the parties paid by these transactions were not reimbursed until the customer final approval, even this would result in no lost funds.

Thus the policy decision as to who will get a refund will have to create some opportunity for fraud for one party (Given the current rates of suspected customer and merchant fraud, allowing the customer leeway appears to be the risk-averse choice.)

This system is money-atomic in that, using collusion, rollback is possible. Merchants can declare bankruptcy and disappear if the billing cycle is too long. Given that a shorter billing cycle means increased costs, this is an economic decision subject to optimization.

5.6.3 Privacy

The information available in an Anonymous Credit Card transaction is shown in Table 5.7.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	None	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf w/warrant	Full	Full	Full	Full	*Full*
Customer Bank	None	Full	Partial	Partial	None
Private Bank	Partial	None	Full	Full	None
Observer	Partial	None	Full	None	None

Table 5.7: Information Available In an Anonymous Credit Card Transaction

Anonymous Credit Card deals explicitly with location information through the communications exchange.

The Anonymous Credit Cards protocol is a high privacy system. Only the Anonymous Credit Card system deals explicitly with location as a pseudonym.

Table 5.7 reflects an assumption that the observer is watching messages to and from the merchant's account. Conversely if the observer were watching messages to and from the customer account then the observer would have partial information about the customer and none about the merchant. It is also possible to use traffic analysis to infer the existence of a transaction between the customer and the merchant's respective addresses. However, there are defenses against traffic analysis, and if this system was adopted the amount of traffic should be sufficient to make traffic analysis problematic.

The merchant bank will know the bank where the merchant's accounts are held, as this information is included in the receipt from the merchant's bank.

The customer's bank does know the total amount spent in a billing period; thus it has an upper bound on any transaction size. Since the billing period is discrete, by definition the customer's bank will know that some transaction occurred within a given time frame. However, exact information on the date and time of transactions, the amounts of specific transactions, and items purchased cannot be read by the customer bank.

Here, law enforcement depends on the record-keeping of the merchant to obtain information about the items purchased.

Of course, any bank or any merchant may guess at the identity of any customer. Using the same pseudonym, or anonymous account, over time can result in a loss of anonymity as other participants update their probability distribution at each pseudonym use. Eventually merchants and pseudonymous banks would become certain of customer identity. This issue has not been addressed, but it appears that it could be solved by changing accounts on a periodic basis.

The Anonymous Credit Card protocol protects critical identity information while allowing law enforcement the ability to access data. Anonymous Credit Cards would provide a high level of privacy.

5.6.4 Regulatory Issues

The Anonymous Credit Card protocol is a technical response to a regulatory problem. Sharing and matching data between financial immediacies is a problem that would negate the advantages of this protocol. Constraints on data sharing by institutions other than credit unions are necessary for this technical approach to work.

Anonymous Credit Card uses information distribution to address the conflict between information availability and privacy. The Anonymous Credit Card protocol prevents merchants and banks from trivially obtaining transaction specific data that can be linked with individual consumers. This distribution of information is provided at the cost of increased complexity and a risk of fraud.

The Anonymous Credit Card protocol provides information for regulatory purposes. Obtaining such information requires a series of court orders, one for each institution.

The Anonymous Credit Card protocol meets both the conflicting needs for information and privacy. However, the Anonymous Credit Card protocol does not provide information for resolution of business disputes in a cost efficient manner.

The Anonymous Credit Card protocol does not necessarily violate banking law in the assumption of a pseudonymous bank, since that entity need not actually hold deposits -- the pseudonymous bank serves as a financial intermediary to transmit financial information. In this way it more closely resembles the psuedonymity server used in the NetBill protocol than a bank.

However, the Anonymous Credit Card protocol assumes that there are controls on information sharing between financial intermediaries. Thus, despite the technical complexity of this protocol, its fundamental value is based on an unsubstantiated policy conclusion. If there were controls on the secondary distribution of financial information, much of the privacy problem that Anonymous Credit Cards proposes to solve would no longer exist.

Finally, the Anonymous Credit Card protocol requires a loss in anonymity every time there is a billing conflict. When there is collusion, presumably all the interested parties get customer and store identity information. Since the protocol is pseudonymous, then the customer identity is exposed to the pseudonymous bank and the customer must change her account. In fact, it appears that Anonymous Credit Cards were optimized for regulatory fitness at the cost of business practicality.

5.7 Summary

First Virtual offers a low privacy and low security system for Internet commerce. First Virtual assumed Internet will be without security, and addressed that lack of security through risk management and loss allocation. Unfortunately, this loss allocation (the merchant losses) limits the goods which are suitable for sale using First Virtual. First Virtual is a low privacy system which further requires merchants to keep extensive records on customers. This reinforces the previous conclusion that the controls created on consumer financial data under the Fair Credit Reporting Act be expanded to cover compilations of non-bank institutions that gather detailed consumer records.

Secure Sockets layer is a first generation Internet commerce protocol that has taken an approach opposite that of First Virtual. First Virtual assumes the Internet is without security and merchant losses are negligible; Secure Sockets Layer assumes that the Internet can be made secure and merchant losses limited by off-line financial management. Secure Sockets Layer does not attempt to provide atomicity, and does not do so. The Secure Sockets Layer is a medium privacy system. The security of the Secure Sockets Layer is limited by the constraints on exporting strong cryptography. This is an argument for the removal of restraints on the export of cryptography.

Secure Sockets Layer may have the risk-creating side effect that many merchants keep records of customer's credit card information on machines connected to the Internet and subject to remote attack. There are at least three possible policy solutions to this problem: security, including cryptography, should be required in popular operating systems; computer operators with inadequate security practices should be liable for all losses caused by their negligence; or data should be deleted as soon as possible, according to the practice recommend in the codes of ethics discussed in Chapter 3.

The Secure Electronic Transactions is a payment protocol which considers all steps in an electronic transaction (recall Section 2.3.2) excluding account acquisition. Secure Electronic Transactions is a low privacy system since purchase information is transmitted in the clear. Secure Electronic Transactions is a high security system as designed, as it removes the opportunities for replay attacks and shared merchant terminals.

The Secure Electronic Transactions standards uses the Secure Sockets Layer for information to be transmitted out of band. Thus the regulatory requirement for weak cryptography has affected the design of electronic commerce systems. This illustrates the ubiquitous effects of constraints on cryptographic exports on electronic commerce and offers an additional argument against removing these constraints.

NetBill provides a high degree of reliability. NetBill can also protect customer privacy in terms of hiding identity information from the merchant. NetBill is not a high privacy system; the central NetBill server keeps records each purchase, including amount of purchase, customer identity, and merchant identity. NetBill has automated a sufficient number of processes that it has the potential to decrease the cost of billing by an order of magnitude.

Anonymous Credit Cards provide both conditional anonymity and reliability. Note that the atomicity, and therefore the reliability, of an Anonymous Credit Card transaction depends upon the funds availability policy of the merchant, customer and pseudonymous banks. Anonymous Credit Cards offers a complex technical solution to a fundamental policy problem, i.e. that institutions share information without consent or control of the subjects of information. Because of the complexity and number of institutions in the system, Anonymous Credit Cards would have a high per-transaction cost.

The analysis of this set of protocols for Internet commerce illustrates that with notational currency, reliability can be simplified by creating a single ledger. The creation of a single ledger means that there is a concentration of information -- thus implying a threat to privacy. The Anonymous Credit Card protocol attempts to address this by accepting the increased complexity as the cost of privacy. However, the relationship between distribution of information and provision of privacy does not always hold true in that increased centralization does not always imply decreased privacy. NetBill, for example, has more centralized transaction processing than Secure Electronic Transactions but provides an equivalent level of customer privacy.

6 Token Currency

"Every one, even the richest and most munificent of men, pays more by check more light-heartedly than he pays little in specie."
Beerhom, 1920

In token currency the strings of bits transferred in a transaction are themselves legitimately valuable. For example, a dollar has value in and of itself and is not a promissory note for a particular transaction from a specific account, like a credit card purchase slip. Because of this transaction independence, token currency need not be linked to specific transactions or identities.

Digicash was the first and remains the canonical token currency system. With Digicash Chaum presented the concept of *blind signatures*. This allows the bank to verify currency for users without being able to identify that currency as it is later spent. Previously any single element token currency provided by the bank to a customer could have been identified by the bank at the time it was deposited. The invention of blind signatures created the possibility for anonymous electronic token currency.

Electronic token currency is particularly interesting in that each new token proposal presents a novel mathematical technique or a novel application of a known technique.

The most difficult problem remains the prevention of double-spending of coins. Any bitstring can be trivially duplicated -- and when a bitstring is itself token currency then there is universal motivation to do so. In the absence of secure hardware, the dominant approaches to solving the problem of double-spending have been limits on anonymity and on-line clearing.

The issue of double-spending is related to the issue of isolation. If a token can be spent more than once then the transactions are not isolated. There is a race condition: whoever gets to the bank first is paid, the second to arrive is unpaid. If there is on-line clearing the payee can clear the token before accepting it. (Clearing refers to the ability of the payee to confirm the validity of a token without depositing it.) In this case clearing enables something akin to two phase commit: clearing is the customer's commitment nested through the merchant, deposit is the merchant's commitment, and acceptance of the deposit is the bank's global commitment. Thus the token is locked until the transaction is complete, so no race condition is possible.

Atomicity is complicated by both the nature of token currency and anonymity. Restoration to the previous consistent state can be difficult in token currency precisely because the token was not necessarily linked to a specific transaction. Anonymous restoration of a previous state is particularly problematic: what if any anonymous individual could lay claim to the money in our wallets?

Consider the three classes of atomicity with token currency. For a token transaction to be money atomic, a customer payment must be linked with the merchant's payment. That is, it must be the case if the customer loses the value of a token then the merchant gains the token; and the merchant gains a token only if the customer loses one.

For a token transaction to be goods atomic, the merchant must only obtain the token if some merchandise was delivered.

For a token transaction to provide certified delivery, the merchant must only obtain the token if the promised merchandise was delivered.

To illustrate the issues of security, reliability and privacy in electronic token currency I begin this section with a discussion of a remote transaction with physical token currency: sending cash through the mail.

I then discuss the original Digicash proposal, followed by a later proposal where anonymity was limited in order to prevent double-spending through the threat of detection. I then consider a proposal for making electronic cash divisible, i.e. making change. Finally, I close with MicroMint which is an electronic token currency with no anonymity.

6.1 Legal Tender: Cash

Cash is token currency. The examination of this case in the section provides a model and basis for comparison with electronic token currency.

In the United States its interoperability is assured by federal law. It is "legal tender for all debts public and private." The business model of cash is interoperability and availability assured by government action to enable and encourage commerce. Internationally, interoperability is provided by currency exchange services, at some price. The consistency in the history of United States legal tender has resulted in global interoperability.

The transaction analysis section discusses the importance of the attributes of cash. This section discusses the problems with trying to use cash to make remote purchases. This is interesting because one fundamental problems remains with electronic cash -- how can a customer prove payment with a remote anonymous purchase? The privacy and security sections illustrate the strengths and weakness of cash along those dimensions.

Cash is interoperable because it is legal tender. The lack of interoperability of bank and state currencies was a driving force behind the creation of a national currency.

There are no limits to scale in the number of users of cash, except those imposed by limits on the number of bills. Not only are individual transactions isolated, but the system is also free from bottlenecks.

The availability of cash has proven critically important for economic and social reasons. The return to the virtual slavery of sharecropping for blacks in the American South was very much predicated by the return to the gold standard and resulting currency shortage. This dramatic example illustrates that there may be unforeseen implications for lack of availability should a single standard for Internet commerce emerge. While the privations of ignorance in an information-intensive society are not comparable to the deprivation of the landless in an agricultural society the delineation between ignorance and interconnection defines the empowered and the powerless in the information age.

Cash is divisible in that there are many denominations; a single high-value token can be exchanged for many low-value tokens, and many smaller tokens can be exchanged for a single high denomination token.

6.1.1 A Transaction

Consider a remote cash purchase -- sending a dollar through the mail to a merchant. It is a three step process. Assume that delivery is to a Post Office Box, so the customer need not offer identity information.

First, the customer gets a dollar from the bank. The bank decrements the customer's account one dollar and provides the dollar to the customer.

The customer then sends the dollar in the mail to the merchant. The merchant verifies the dollar through visual examination -- this is analogous to off-line verification. The merchant can prove that he has the right to spend that dollar by virtue of having the dollar. Thus the merchant has no need to provide authentication to spend the dollar.

When the merchant deposits the dollar in he bank, the bank will not link that dollar to the one given previously. In theory the bank could keep track of the serial numbers of all the dollars it gives out and to whom it gives these dollars, but this would be an extremely costly method of surveillance.

The customer cannot prove that the merchant received payment; this is not an atomic transaction. The merchant can simply take the money to the bank for deposit. The customer cannot prove previous ownership of the dollar or commitment to deliver merchandise on the part of the merchant.

Consider the case if the bank kept records of the identities of all those who withdraw dollars linked to the serial number of the dollars that were withdrawn, and dollars could only be spent once before being returned to the bank. Then by sacrificing anonymity the customer could claim to have paid the merchant by knowledge of the merchant's identity and the serial number. Of course, in this case one could simply guess. This case illustrates a conflict between anonymity and atomicity present in the electronic systems examined in this chapter.

This transaction is isolated. Regardless of what occurs in other transactions, the merchant can deposit the dollar.

This transaction may not be consistent. If the dollar is lost in the mail then the customer believes that she has been defrauded and the merchant does not know a transaction was attempted. The dollar may simply disappear.

The transaction is durable. The merchant will have the dollar; the customer will not. The customer cannot arbitrarily reverse the transaction.

6.1.2 Security

Cash does not require trust between users. If a bill is determined to be counterfeit, the holder of the bill is not compensated. The validity of a bill can be partially verified during the transaction by visual inspection. By accepting cash, merchants imply only that they trust their own ability to detect counterfeit.

Clearly there are security failures in the form of counterfeit notes, but security is generally maintained by a time-tested work factor. The design of the bills is periodically updated to discourage counterfeiting. Systems-level failures in the paper currency system are prevented by risk-limiting regulation, federal depository insurance, limits on denominations, and the sheer magnitude of the task of passing enough counterfeit currency to upset the entire system.

However, once the hurdle of printing a single counterfeit bill is met then the marginal cost of creating another dollar approaches zero.

The dollar could be taken from the US mail since it is unprotected. In 1994 there were 20,976 million pieces of mail delivered (Bureau of Census, 1995). Thus the sheer magnitude of searching the mail, combined with the relative rarity of finding cash in such an endeavor, provides a high work factor for the prevention of theft by observers. There is no advantage to scale in this sort of theft: searching for the nth dollar will be as hard as searching for the first.

6.1.3 Privacy

Cash offers both privacy and anonymity because a dollar does not contain information that can be used to determine its transaction history. Neither does the exchange of cash necessarily create a record including the identities of those involved. Cash transactions usually provide anonymity of the buyer but not the seller. The privacy of cash is limited by the potential for physical observation. The information available to different parties in a cash transaction is shown in Table 6.1.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	Partial	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf	None	None	None	Partial	None
Bank	None	None	None	Partial	None
Observer	Full	Full	Full	None	None

Table 6.1: Information Available In a Cash Transaction

There are no bank or law enforcement records produced in a cash transaction. It is reasonable to assume that no bank employee or law enforcement officer observes most cash transactions. Therefore the information available to a bank or to law enforcement is limited by what they would obtain from written records. Reporting of some transactions is required by law, but these reports depend on the active cooperation of the parties involved. The bank has an upper limit on the size of any transaction, since the bank knows the amount of any resulting deposit.

In a remote transaction the customer can choose to have materials delivered to a Post Office Box, so only the customer's region of residence is known.

Here we consider an observer who is physically well-placed: for example the observer is beside the customer in the post office. The observer can watch the item be placed in the post, but cannot open the envelope. Again, the work factor makes it unlikely that the letter will be intercepted, although it may certainly be lost.

6.1.4 Regulatory Issues

The current regulatory structure was built over time with the assumption of paper money. The fundamental assumption in this regulatory structure is that risk is placed on the person most likely to be able to prevent loss. For example, if the merchant steals the money the customer absorbs the resulting loss. This is because the customer is the only person empowered to choose to send her cash in the mail. Similarly, merchants and banks lose if they accept counterfeit cash because merchant and banks are in the best position to prevent counterfeiting.

This principle has not yet been applied to electronic commerce. This is partially because the ability and responsibility in terms of keeping information secure has not yet been culturally determined, and partially because there has been no market failure thus far to allocate risk in a way that is acceptable to consumers.

6.2 Digicash

In Digicash (Chaum, 1985) the customers hold the monetary value in the form of electronic tokens. Customers and merchants exchange tokens. These tokens are validated by the bank.

Digicash provides a mechanism for electronic payment. Digicash protocols do not provide mechanisms for discovery, negotiation, delivery or conflict resolution. The scope of Digicash is both its strength and weakness. The advantage is that Digicash can provide an elegant and simple protocol. The disadvantage is that Digicash cannot offer to decrease the cost associated with collection and dispute resolution. In fact, Digicash is specifically designed to mimic cash so that only the purchase itself and the detection of counterfeits are properly the business of Digicash.

6.2.1 A Transaction

The steps in a Digicash purchase are shown in Figure 6.1. This protocol was the first use of blinded tokens for electronic cash. Recall that any party *X* has asymmetric public keys with public key X and secret key x.

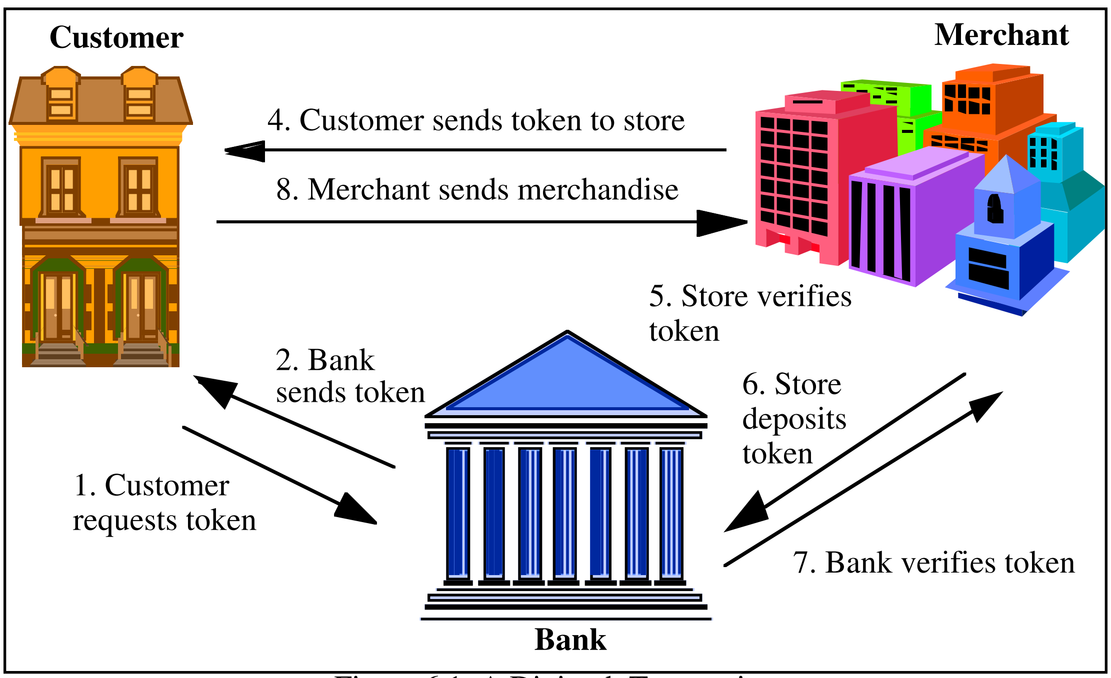

Figure 6.1: A Digicash Transaction

The customer selects a random number and constructs a token. Then the customer encrypts the random number with the public key of the bank's asymmetric key and multiplies it with the token and send it to the bank for validations: $r^B t$. The bank signs the token with its corresponding secret key, and returns: $(r^B t)^b = (r^B)^b t^b$. The customer then divides by the random and gets a token: $(r^B)^b t^b)/r = (rt^b)/r = t^b$. This token has been validated by the bank and has been recognized by the banks as valid. However, the bank has never seen the token and could not distinguish it as the token given to the customer.

Thus the exchange is as illustrated below. All calculations are done modulo some prime p.

1: C -> B	$r_B t$	(mod p)
2: B -> C	$(r_B t)_b$	(mod p)
3: C -> M	t_b	(mod p)
4: M -> B	t_b	(mod p)

Notice that the explicit description provided in the original description of Digicash ends with verification of the token by the bank for the merchant. It is exactly this definition of a transaction that causes problems in terms of atomicity.

Notice that the token received by the bank in step five cannot be identified as the same token sent out in step two. This is the critical element that makes Digicash anonymous.

The information necessary for conflict resolution or dispute prevention is not provided by Digicash.

In fact, if the protocol is interrupted between step four and the delivery of goods to the customer, then the customer has effectively been defrauded. Since the customer is anonymous, he or she cannot simply contact the merchant and ask for the goods to be resent. (The loss of anonymity possible with location information has a positive effect here in that merchants could send a second time to the same IP address.) The merchant could also claim not to have received a token while cashing the token in at the bank. In this case the customer is again defrauded. In no case does the customer have any basis for complaint other than her own testimony.

6.2.2 Security

Digicash assumes the privacy of cryptographic keys: the bank's, the customer's and the merchant's. Consider the results if one of these assumptions is invalid.

An adversary who gains access to a bank's private key can generate counterfeit tokens that are indistinguishable from valid tokens. These tokens can be generated in any amount desired, so compromise of the key compromises all tokens in circulation.

An adversary who gains access to a customer's private key can drain a customer's account. This could be mitigated by including a challenge and response series, where the bank keeps only the hash values of answers for a set of questions (This technique is used in anonymous credit cards, as discussed further in Section 5.6.) Since digitally signing a token is four orders of magnitude more processor intensive than verifying a hash value, this appears to be an reasonable addition to total processor load.

Access to the merchant's private key could either allow the attacker to defraud customers, or empty the merchants account. If customers encrypt tokens with the merchant's private key to prevent theft and assure privacy, then the attacker can obtain tokens through eavesdropping. Similarly, the attacker could imitate the merchant, accepting orders and payment. If the merchant's bank accounts are protected only with public keys then the merchant's account may be drained as well.

Two tokens can be multiplied to construct a third, as follows: $(n_1)b\ (n_2)^b = (n_1\ n_2)^b$. Notice the counterfeiter still has possession of the original tokens. Tokens can be multiplied to form new valid tokens; that is, consumers and merchants can trivially manufacture cash.

Digicash transfers are not money-atomic (Yee, 1994). The customer may attempt to resolve this state by canceling the token (by cashing it in), but if the merchant also does this, the result is a race condition (This also violates consistency and isolation.) Since Digicash cannot reveal who cashed in a token when a merchant claims that he did not cash in the token, and the customer claims that the merchant did, dispute resolution can be a problem.

Another option for addressing issues of customer double-spending and merchant fraud is to assume the customer is always right. This would require only keeping the names of customers that complain and the merchants that are the subjects of their complaints. This would maintain the anonymity in a successful transaction of Digicash while reducing the risk of consumers. An aggressive technique of disallowing merchants suspected of fraud may limit the popularity of the system since consumers and merchants are drawn to popular systems.

A second option is to assume that the customer is always the fraudulent party. This is the option chosen in the Digicash alternative where the detection of double spending is enabled through embedding identity information in the token. This protocol is examined in the next section.

6.2.3 Privacy

The information available to the parties in a Digicash transaction is shown in Table 6.2. Recall that partial identity information trapped by the observer and the seller result from location information.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	Partial	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf w/warrant	Full	None	None	None	None
Bank	Full	None	Full	Full	None
Electronic Observer	Partial	Partial	Full	None	None

Table 6.2: Information Available In a Digicash Transaction

Digicash is a high privacy system. The merchant has only the information necessary to assure payment, and the bank has only the information necessary to credit or debit an account.

6.2.4 Regulatory Issues

Digicash does not provide information to law enforcement. This implies that Digicash would be an excellent instrument for money laundering or other illicit purchases. However, this is mitigated by the fact that tokens must be verified on-line, and therefore banks can identify large deposits or transfers. That the Know Your Customer regulations (31 CFR §103) would not be affected by the type of currency a customer deposits will serve to assure that bank transactions remain accessible for auditing. This suggest that limits on anonymous account transfers apply to Digicash, as they do to analog cash. That Mark Twain is offering Digicash accounts supports the conclusion that regulators are willing to accept anonymous currency if it enters and exits the electronic realm through auditable channels.

The unlimited liability of a customer with the loss of a private key may violate the Electronic Funds Transfer Act. Thus any cost of fraud is transferred to the customer. This business assumption may not be valid, especially in the United States, due to the Electronic Funds Transfer Act. The Electronic Funds Transfer Acts specifically limits consumer loss in electronic funds transfers to $50 per lost instrument. It is not certain if a Digicash account meets the definition of an instrument.

The lack of any receipt and the ease of merchant fraud seem to create problems with the Truth in Lending Act and Electronic Funds Transfer Act requirements for receipts and billing[20], as implemented in Regulations E and Z, respectively. There is a technique to provide receipts and certification of merchant's commitment in an anonymous system, as shown in Chapter 6. However, this technique significantly adds to the complexity of anonymous exchange of digital tokens. Furthermore it significantly extends the scope of the transaction beyond that currently considered by Digicash.

Digicash-based banks can provide aggregate information, and are certainly capable of storing records on individual withdrawals and deposits. The anonymity of Digicash means that the bank cannot link the deposits to the withdrawals. It also means that coins cannot be traced through their path if they change hands more than once. However, transferring a coin more than once creates the risk that previous possessors of the coin will return it before the subsequent owners.

6.3 Digicash with Detectable Double Spending

In the earliest version Digicash offered an anonymous protocol graceful in its simplicity (Chaum, 1985). As shown above, the protocol was not secure for on-line commerce.

The later version of Digicash (Chaum, 1989) considered here addressed two of these problems. Tokens can no longer be trivially manufactured. Individuals that double spend can be detected, with some probability, after the fact. Probability of detection is a function of the size of the fraudulent purchase, and is based on the assumption that only the account holder will double spend. Consumers can verify that they have paid merchants; however, this verification requires the loss of anonymity.

One alternative form of Digicash (Brickell, Gemmell and Kravitz, 1995) with preventable double spending attempts to maximize socially important data while reducing the threat of data surveillance by allowing customers to withdraw a certain amount anonymously while larger amounts are tagged for later information. However, tagging of the information requires the use of trusted hardware with an observer built in. This latter scheme depends on the existence of secure tamper-proof hardware trusted by both the consumer and the government in the consumer's smart card.

Other forms of digital cash have addressed the problem of divisibility: that is, the problem of making change (Brickell, Gemmell and Kravitz, 1995;Okamoto and Ohta, 1991). One version offered pseudonymous coins (Okamoto and Ohta, 1991), meaning individual coins could be anonymous but segments of the same coins could be linked (Brickell, Gemmell and Kravitz, 1995).

---

[20] It is worth noting that the revision of Regulation E noted in Chapter 3 would exempt stored-value cards with storage capacity of less than $100 from the receipt requirements. Digicash is suitable for smart card implementations, yet my focus here is upon Internet commerce.

6.3.1 A Transaction

Auditable Digicash transactions require three parties: a bank, a merchant and a customer. The steps in a Digicash transaction with identification of double-spenders is shown in Figure 6.2.

In the first step the customer formats a series of potential tokens. This requires two well known one way functions, g and f. All operations are modulo n, where the bank knows the factorization of n. The form of the token is

r_i^3, $f(x_i, y_i)$

where

$x_i = g(a_i, c_i)$ and $y_i = g(a_i \text{ XOR } (\text{account number} \,\|\, (\text{counter} + i), d_i)$

Here $\|$ refers to concatenation and r_i is a random number. The account number and counter are known to both the bank and the customer.

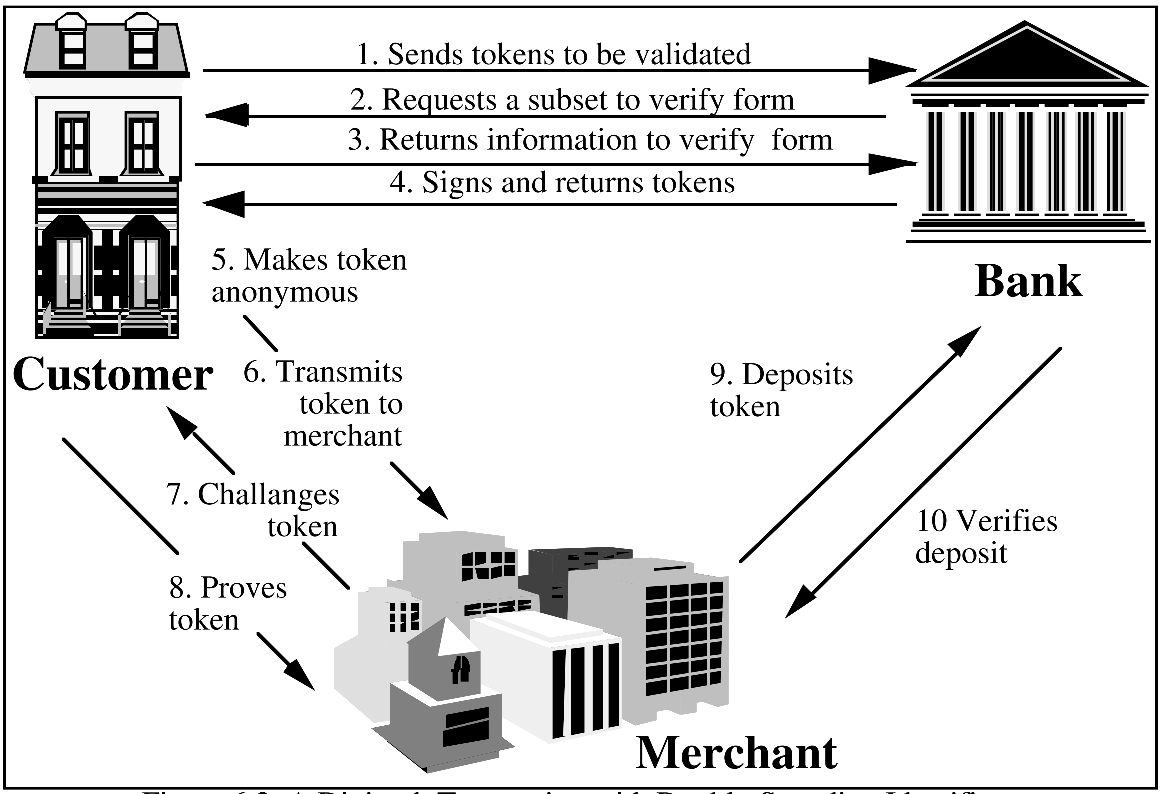

Figure 6.2: A Digicash Transaction with Double-Spending Identifiers

In the first step the customer formats a series of potential tokens. This requires two well known one way functions, g and f. All operations are modulo n, where the bank knows the factorization of n. The form of the token is

r_i^3, $f(x_i, y_i)$

where

$x_i = g(a_i, c_i)$ and $y_i = g(a_i \text{ XOR } (\text{account number} \,\|\, (\text{counter} + i), d_i)$

Here $\|$ refers to concatenation and r_i is a random number. The account number and counter are known to both the bank and the customer.

Thus in the first step the customer authenticates themselves to the bank through the use of the customer's public key. The customer sends a set of tokens, the construction of which was described in the preceding paragraphs.

In the second step the bank challenges the customer to show that the tokens were constructed properly. This means after the customer sends a number of tokens to the bank, the bank selects a subset of these and returns them to the customer.

In step three the customer must return the elements of the token to the bank (the appropriate r, a, c and d) and the bank checks to see that the challenged tokens were of the proper form. The bank then discards these tokens, and signs the remaining blinded tokens. The random selection of a subset of presented information for verification, while allowing the other information to remain hidden, is called *cut and choose*. Of course, there is a chance that the customer can obtain a signature on a problematic token which will enable the customer to commit fraud. The bank can determine its willingness to accept the risk of fraud and implement the cut and choose technique in a manner consistent with that level of risk aversion.

In step four the bank returns the signed blinded tokens.

In step five the customer divides by the random number r_i, thereby obtaining a properly signed, but anonymous, token. The consumer and the bank then increase their corresponding counter values appropriately. The analog of this is that the customer has less money in her account but more money in her virtual wallet.

In step six, the customer selects an item which costs a number of tokens and transmits the selected tokens to the merchant.

In step seven the merchant challenges the tokens. Unlike the bank, the merchant challenges all the tokens. However the customer provides only part of the information necessary to verify customer identity. For each token, the merchant can request either the appropriate a, c, and y values, or the appropriate x, (a XOR (account number || (counter + i)), and d values. Either set of these values allows the merchant to verify the token. However, with both of these the merchant could deconstruct the token and identify the customer through her account number.

In step nine the merchant sends the token and the values used for verification to the bank for credit.

In step ten the bank credits the merchant and verifies that the tokens were not double-spent.

For every token that the customer spends there is a 50% chance that the next merchant will request a different set of values for token validation. If a different merchant asks for a different set of verification values, then the bank will have enough information to identify the account and therefore the account holder. Thus, double spending is limited by the ability of the bank to detect the fraud, and the correspondingly high penalties. Note that the knowledge of the token authenticates the customer as being authorized to provide payment.

6.3.2 Security

Digicash transactions have high privacy and potentially low transactions cost. However, Digicash transactions may be subject to a high fraud rate.

Digicash fails to fully address merchant fraud. If a merchant receives a token and then deposits it, the merchant can claim not to have received the token. The customer can provide the corresponding a, c and d values and thus illustrate to the bank that it is indeed the customer's account number embedded in the token. This means that a customer can prove payment at the cost of loss of privacy. Thus money-atomicity is possible with this version of Digicash, but it is at the cost of anonymity. However, this possibility is not included in the protocol. In fact, if the merchant claims to have lost the token and the customer spends it again, the customer herself is at risk for fraud prosecution. Similarly, if a customer loses Digicash tokens and the embedded information it is unlikely that the thief cares if the account owner is identified as a result of the thief's double spending.

However, Digicash creates no record of any sales agreement or delivery between the merchant and the customer. This means that the customer cannot prove that fraud occurred. Thus there is neither goods atomicity nor certified delivery.

Most items will cost some combination of tokens. Consider the last time you paid for an item with a single denomination and received no change. However, if there is a case where an item costs exactly one token, attempted fraud would have a 50% chance of success.

In this Digicash protocol it is always assumed that the customer is the dishonest party. Since most credit card fraud results from unauthorized use of cards through theft or loss where the owner of the account cannot prevent the fraud, the validity of this assumption is questionable. If the customer is indeed committing the fraud, the assumption that detection after payment is sufficient to reduce risk is also questionable, given the opportunity to recycle funds and disappear after successful fraud occurs (McClellan, 1995). The same detection mechanism has failed to deter checking fraud.

The results of a loss of merchant or server security remain unchanged from original Digicash. Thus there is the risk of long term undetected subversion of a Digicash server.

6.3.3 Privacy

As shown in Table 6.3, the information available to various parties in a valid transaction is the same with Digicash with and without detection of double spenders. In the case of a fraudulent transaction the identity of the original account holder can be determined.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	Partial	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf w/warrant	Full	None	Full	Full	None
Bank	Full	None	Full	Full	None
Electronic Observer	Partial	Partial	Full	None	None

Table 6.3: Information Available to the Parties In a Digicash Transaction

Thus, Digicash with detection of double spenders remains a high privacy system.

6.3.4 Regulatory Issues

The ability to identify double spenders does not affect the overall availability for law enforcement, as shown in Table 6.2. Thus the problems in terms of providing information that existed with the original Digicash remain.

An additional problem is that the assumption that an account holder is always at fault may not be legally valid. If an attacker can obtain the information to steal a token, the attacker can spend that token many times. Thus an account holder may have very high liability in the case of lost tokens. This unlimited liability seems to violate the Electronic Funds Transfer Act.

Again the Truth in Lending Act requires that customers be able to dispute purchases. In this case, the customer can dispute purchases at the cost of anonymity. Thus this is a significant improvement in terms of regulatory fitness over the initial version of Digicash.

6.4 MicroMint

MicroMint and PayWord are a set of electronic commerce protocols (Rivest and Shamir, 1996) that use the difficulty of calculating hash values and the birthday paradox to provide electronic currency. Despite their mathematical similarities, PayWord is a notational, credit-based scheme and MicroMint is a token, debit-based scheme.

MicroMint is problematic because it assumes the solution to the problem of electronic commerce. There is a bootstrap problem. Once the coins are established and the consumers have coins, the customer purchases coins and merchant redeems them. However, the distribution of the first generation coins remains unspecified.

MicroMint calls the bank a “broker”. Instead of holding deposits, the broker generates coins and exchanges them with customers. The broker has the float until the coin is spent. In order to defeat double-spending the broker must know the customer identity. This knowledge requirement, the consumer’s need to be able to draw coins, and the merchant’s ability to deposit them all suggest that it is a reasonable conjecture that the broker is a bank in all but semantic terms.

The bank mints coins by using k hash values and the birthday paradox, as discussed in the second chapter. Rivest and Shamir have calculated that to obtain one k-way collision requires calculating expected $2^{(k-1/k)}$ hash values; however, to get c k-way collisions requires expected $c2^{(k-1/k)}$ hash values. The authors compare this to the initial investment in a minting facility, with high initial costs and low marginal cost for each additional bill printed. A coin is a set of numbers that all have the same hash value, thus a coin is a set of k numbers $(x_1,x_2,x_3,...x_k)$ where $h(x_1)=h(x_2)=h(x_3)=...=h(x_k)$.

Using the publicly known hash values merchants can verify off-line that currency sent by customers is of the valid form. In order to prevent double spending, the customer’s identity is embedded in the hash values sold to a customer, like so:
$h(coin) = h(x_1,x_2,x_3,...x_k) = h(\text{customer identity})$.

By using a lower hash value this becomes computationally feasible. A 16-bit hash value is recommended in the original publication.

The cost of generating coins is shown in the table below, as excerpted from (Rivest and Shamir, 1996).

Table 6.4 clearly illustrates that small scale attacks are not feasible.

This is for the case where a coin consist of four numbers, i.e. a coin requires four collisions. The numbers are 36-bits. Thus there is a tremendous initial investment in preparing the first coin, but as the number of coins created increases, the cost decreases.

Number of Hashes	Coins Produced	Hashes/coin
$2^0 \ldots 2^{26}$	0	•
2^{27}	1	2^{27}
2^{29}	2^8	2^{21}
2^{32}	2^{20}	2^{12}
2^{36}	2^{32}	2^3

Table 6.4: Returns to Scale in Minting Money through Hash Collisions

Large scale attacks are mitigated by changing the hash value[21] monthly. All forged coins are invalid at the beginning of the month, and the forger cannot begin generating hash values until the hash function for the month has been announced. Furthermore, the broker can detect forged coins, announce a new hash function at any time, and use hidden predicates for daily updating. [A hidden predicate is a special characteristic of a number which is not apparent upon simple examination. To generate a number with hidden predicates, some of the bits in the number are made a function of other, random bits.]

6.4.1 A Transaction

MicroMint transaction begins when the customer obtains a coin from the broker. The steps in a MicroMint transaction are shown in Figure 6.3 below.

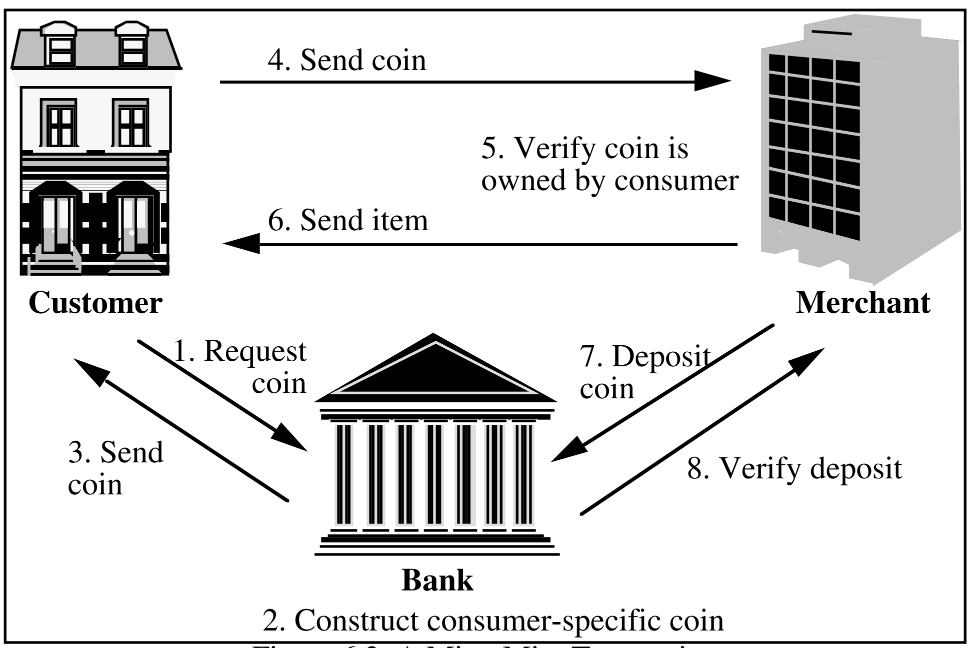

Figure 6.3: A MicroMint Transaction

In step one the customer requests a coin. This customer must authenticate her identity in order to obtain the coin. Customer authentication for purposes of coin generations is not specified. Thus the information exchanged at that step cannot be determined. Note that merchant authentication and broker authentication are similarly unspecified.

In step two the bank must have the customer's identity in order to properly construct the coin. The bank then constructs a provably valid coin linked to the customer's identity as

[21]The authors suggest using the same hash value with a different salt.

described above. Notice the bank has an extremely large database of known hashes from which to construct a coin -- the bank does not need to begin anew at every customer request.

In step three the bank delivers the requested tokens. After the third step the customer can prove that she has a valid token, and can prove her ownership of that token. Proving ownership requires exposing her identity.

In the fourth step the customer transmits the item, price, token(s), and her identity to the merchant. The merchant will also have the information transmitted during browsing at this time. The customer's identity verifies her ownership of the coin. However, the mechanism of customer authentication is not addressed in this protocol.

In the fifth step the merchant verifies the token and the customer’s ownership.

In the sixth step the merchant delivers the item. This is not explicit in the protocol; however, it is assumed to be part of the protocol since the protocol is off-line. Thus the deposits are presumably batched.

In the seventh step the merchant verifies that the token has not been spent.

Like the later version of Digicash, this assumes that only users double spend. The entire issue of authentication is left unconsidered. Presumably the customer would have a certificate or registered public key to authenticate herself to the merchant or the broker.

The delivery of merchandise is not included in the protocol. Thus the protocol cannot be money atomic or goods atomic.

Off-line transactions are not isolated. If a coin has been previously spent, the merchant is not reimbursed. Thus the outcome of one transaction depends on the existence of another.

On-line transactions are consistent. Since MicroMint is not anonymous, the customer can inquire as to whether a merchant has deposited that customer’s coins without the broker allowing anonymous inquiries into merchant records. The merchant can verify a coin before accepting it. The customer may not have gotten any merchandise, but the money transfer will be consistent. Off-line transactions are not consistent, because a merchant may believe that a coin has not been previously spent while the customer knows she is committing fraud.

The transactions are durable after clearance with the broker but not before. If a customer spends coins in two locations there is a race condition. The merchant paid may not be the merchant that is credited at the broker. Thus off-line transactions are not durable.

6.4.2 Security

Security parameters include the strength of the hash value, the secrecy of a hash value before it is released, and the number of collisions required for a coin.

If the broker does in fact hold deposits then the issue of customer authentication to the broker needs to be addressed. MicroMint recommends that the broker have a shared DES key with each customer. This could be used for authentication as well.

If the hash value chosen for the next month is leaked to an attacker, then the attacker can create coins as quickly as the broker. Since the attacker has lower costs than a legitimate broker presumably the attacker can invest as much as the broker in processing power.

If the predicates are used then the attacker would need both the predicates and the hash parameters to commit forgery. Thus the use of predicates addresses these security issues in a cost-effective way.

An increase in the number of collisions required for a coin increases the cost of coins to the broker and the cost of verification to the merchant. Conversely, increasing the number of collisions required increases the security of the coins. In this way the number of collisions required is a parallel to certificate lifetime, as discussed in Chapter 2.

Simple observation of other systems suggests that some proposed security measures against large scale fraud may be ineffective. The broker can recognize false coins, just as in the physical world the bank can recognize bad checks. This has not proven effective with check fraud, precisely because checks are verified at the bank, not at the merchant. The merchant takes the risk while the broker has the ability to detect and prevent the fraud.

The broker can also declare a current hash period to be over and recall all coins. Consider that attackers can invest in computing power equal to the broker, have no customer overhead and obtain all goods purchased with false coins for no cost. Consumers may pay for discount coins and the attacker does not have to reimburse merchants for services.

However, the use of daily predicates, the ability of the broker to select the hash function and the computational overhead to produce the first coin all provide strong barriers to potential attackers.

6.4.3 Privacy

The information available to various parties in the MicroMint system is shown in Table 6.5.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	Full	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf w/warrant	Full	Full	Full	*Full*	Full
Bank/ Broker	Full	Full	Full	None	Full
Electronic Observer	Full	Full	Full	Full	Full

Table 6.5: Information Available In a MicroMint Transaction

Here again the ability of law enforcement to obtain information about purchases depends on the record-keeping of the merchant.

MicroMint is a low privacy system. Since the creator of coins is modeled as simply a broker, the ability to provide pseudonymous services is limited. If the broker were in fact an account holder for the various consumers then the broker could easily offer pseudonymous coins. The cost of psuedonymity would be one search of the consumer database. Consumers could change pseudonyms when no coins were held under a pseudonym. Because of re-spending the broker would store pseudonyms until the hash value for the pseudonymously released coins was invalidated.

Consumers can only spend their own coins. This means that there is no threat of security loss if someone copies your coins during a transaction. Because of this security, there is no extension to the protocol for encrypting negotiation and payment. Thus an observer can obtain all transactional information about the purchase. This protocol could be combined with any product or protocol which provides encrypted peer to peer communication on the Internet.

By offering pseudonyms and protecting merchant to consumer transaction the information matrix for MicroMint would be changed as shown in Table 6.6.

Changes are shown in boldface. This would make MicroMint a medium privacy system.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	**Partial**	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Law Enf w/warrant	Full	Full	Full	*Full*	Full
Bank/ Broker	Full	Full	Full	None	Full
Electronic Observer	**Partial**	**Partial**	**None**	Full	**None**

Table 6.6: Information Available in an Enhanced MicroMint Transaction

Notice that offering pseudonyms would require some changes in the authentication. The broker would need to either sign pseudonymous keys or provide pseudonymous certificates. Presumably the latter is preferable because the protocol could then remain off-line from the perspective of the merchant.

6.4.4 Regulatory Issues

MicroMint offers inexpensive transactions, at the cost of anonymity and individual security.

MicroMint can fulfill all the requirements for information for regulatory purposes.

Individual security is lost in that there is no protection against malicious framing. The argument is that "the known mechanisms for protecting against such behavior are too cumbersome for a light-weight payment system." Given the amount of motivation individuals have to harm one another, as clearly illustrated in the records of law enforcement in every community, this is a significant hazard.

Effectively, the lack of protection from malicious framing prevents the use of this system by anyone in a high-profile position, those who publicly hold controversial views, those who support unpopular causes and even those in contentious relationships. Given that one can be the object of derision or harassment for gender, race, or orientation, this is arguably a commerce system built for heterosexual white men who are quiet about any unpopular beliefs. Given that such a design decision may effectively exclude classes of people protected under the Equal Credit Opportunity Act, should there be regulatory action defining the minimal requirements for electronic commerce systems? Should these minimal requirements include protection against malicious framing?

This case, where malicious framing is not considered, represents a clear case where the risks are taken by customers and merchants to save effort on the part of the bank. This is not a system that appears to be prohibited by current regulation: the consumer cannot lose more than $50 for a lost instrument. Yet this is a system which clearly violates a basic principle of public policy: the risks of loss should fall upon the party most able to prevent those losses. The merchants can lose money; the customers can lose commerce privileges; yet only the broker can prevent such losses.

In this case, there is a clear policy principle at stake. Yet the market has not yet failed to address this issue, so preemptive regulation may be unnecessary. Presumably the broker is actually a bank, since there must be some deposits against which the customer draws. Either that, or this system is interoperable with other electronic commerce systems. If it is the former, regulators can examine the system and decide if it is acceptable. If it is the latter, then other providers of electronic commerce will decide. If the other provider of electronic commerce provides the user the ability to contest charges, the market may in fact push the final cost of lost money on the broker, as customers object to denial of service or charges for stolen coins.

6.5 Summary

Digicash is graceful in its simplicity and offers complete anonymity to the customer. Yet Digicash offers complete privacy at the cost of low reliability. Digicash offers neither money nor goods atomicity.

In the later version of Digicash addressed in Section 6.3, Chaum attempted to prevent double spending, thereby increasing reliability, through encoding identity into each token to be spent. Encoding identity allows double-senders to be identified, thereby resolving the conflict between anonymity and accountability in the case of double spending. The addition of integrity provides sufficient information for dispute resolution in issues of payment, but not enough information to resolve disputes over goods delivery.

MicroMint has the potential to create anonymous currency economically for a large number of users. By creating digital currency using a process with decreasing marginal cost, MicroMint can provide anonymous token currency to a large number of consumers. MicroMint would be economical for micro-transactions, which are too small for billing or collection using current techniques.

MicroMint in its most simple form offers no money atomicity. In order to provide money atomicity MicroMint is extended so that customer identity is included in every coin. Thus the extension of MicroMint to preclude double spending depends on the requirement that every consumer identify herself to the merchant to verify the her right to spend a coin.

Micromint, along with the two versions of Digicash, illustrates the trade-off between atomicity and anonymity. Currently with electronic token currency without trusted hardware, there are explicit trade-offs between privacy and atomicity.

7 An Anonymous Certified Delivery Layer

"A disease known is half cured."
Fuller, 1732

Can transactions using anonymous electronic currency be atomic? Or is the cost of anonymity exposure to fraud? The relationship between anonymity and atomicity in electronic transactions has been an open question (Tygar, 1996). In fact, anonymous transactions can have the highest level of atomicity. In this chapter I demonstrate this assertion by example.

In this work[22] I present a protocol for anonymous certified delivery that removes the need for any trade-off between atomicity and anonymity. Counter-intuitively, anonymity is provided through the creation of a publicly readable transaction log. The transaction log provides forced serialization; in fact, this system is expanded to use two phase commit. This system is optimized for information goods delivered on-line. While it can also be used to provide detailed receipts for other types of purchases, the strength of certified delivery that guarantees delivery cannot be carried into the physical realm.

The anonymous certified delivery protocol can be used with any on-line anonymous or pseudonymous token currency.

We assume all parties can perform basic cryptographic operations, e.g. cryptographic hash computation, signature computation and verification. All parties have well-known or verifiable public keys. All signatures can be verified by the receiving party.

Other assumptions are that the secret elements of public keys are not disclosed; that tokens are verifiable or can otherwise be trusted; and that the bank can distinguish those tokens to be used with certified delivery. Distinguishing tokens to be used with certified delivery enables the bank to refuse their deposit without appropriate transactional records.

To provide atomicity, a modified, cryptographic version of the two-phase commit protocol is implemented using an external, publicly accessible transaction log. The transaction log receives and records messages and then reproduces the recorded messages. The log localizes the global commit decision to a single entity. The log also acts as a time-keeper, aborting transactions that expire

Communication channels are assumed to be secure. The method used to implement this security (e.g. public keys) is not important to the functioning of the protocol. Some messages could be left unsecured, improving efficiency at the cost of a limited loss of privacy, but the details of this procedure are not currently addressed.

As with NetBill and SET, message signatures are computed on hash values of the plaintext and then appended to the plaintext to form a signed message.

The transaction log and bank are separated only for clarity of exposition. There is no reason that the entities referred to here as transaction log and bank are not merely separate machines in the same facility, analogous to the ledgers and vault of a depository institution.

Each transaction has an expiration time. This implies that the banks and the logs have synchronized clocks. The customers and merchants are then responsible for synchronizing with their log or bank of choice. Although this is a nontrivial problem, there are established approaches to the problem of synchronizing distributed secure transactions (Smith, 1992).

[22]An alternative discussion of this protocol is available in (Camp, Harkavy, Tygar and Yee, 1996)

7.1 A Transaction: Preparation & Purchase

Figure 7.1 below illustrates the preparation steps required in an anonymous atomic purchase.

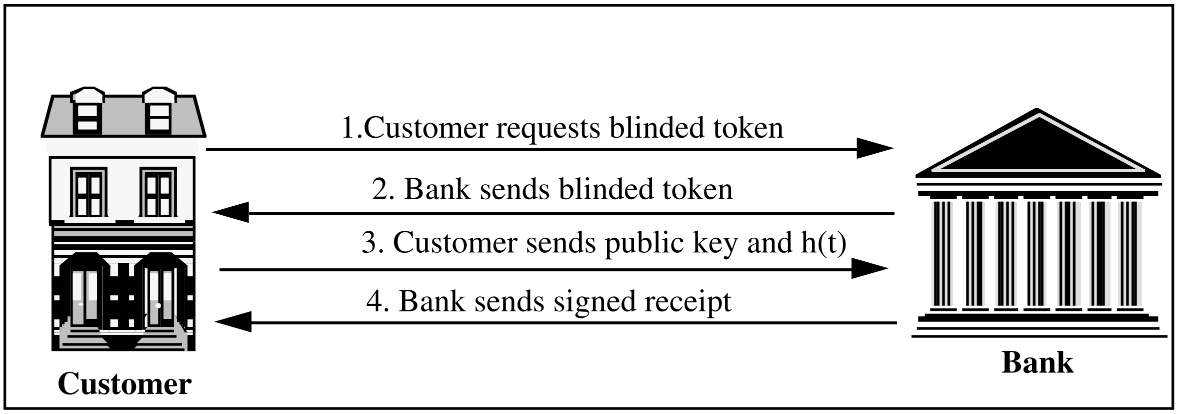

Figure 7.1: Preparation for an Anonymous Atomic Transaction

In the first step the customer obtains a token from her bank. Since this layer could be an addition to many anonymous token protocols, the technique for generating the token and the form of the token are not specified. The only specification necessary is that the resulting token be anonymous and, as noted above, identifiable by the bank as for use only with certified delivery.

We assume the customer generates a public private key pair to be used uniquely for the purpose of assuring certified delivery for this particular transaction. Call this pair q and Q to distinguish them from the pair c and C normally used by the customer and linked to her identity. The key pair qQ is linked to the ownership of the token, not to the identity of the customer. (Note that I do not address the issue of traffic analysis in this protocol.)

Thus the steps are:

3: C -> B	h(anonymous token),Q
4: B -> C	$(h(\text{anonymous token}),Q)_b$

Key pairs q and Q can be generated off-line in advance of purchases. Notice that the key needs only to be strong enough to make obtaining the token cost prohibitive.

In the fourth step the bank returns a signed receipt for the token and the public key. This receipt enables the customer to demand that the subject token not be accepted for deposit unless accompanied by a receipt signed by the token specific key, q, indicating customer commitment to the transaction. The receipt is, in effect, a public key certificate, issued by the bank, binding the public key Q to the rightful holder of the token t.

The customer may now begin the transaction, as shown below in Figure 7.2. The steps are numbered consecutively, with the assumption that they follow the preparation steps above. Each step also has an associated label to simplify the following discussion of atomicity. The customer is assumed to have the merchant's certificate and key; and the merchant is assumed to have the certificate that binds the public key Q to the rightful holder of the token. As with previous protocols, the exchange of certificates is not shown.

browsing	P1: C -> M	discovery & negotiation (not specified)
contract	P2: M -> C	$(\text{contract}, E_k(\text{goods}), \text{TID}, L, Q)_m$
customer commitment	P3: C -> B	$(t, \text{expiration}, M, L, \text{price}, \text{TID}, Q)_q$
notification	P4: B -> M	$(\text{expiration}, M, L, \text{price}, \text{TID})_b$
merchant commitment	P5: M -> L	$(\text{expiration}, k, \text{TID})_m$
global commitment	P6: L	$((\text{expiration}, k, \text{TID})_m)_l$
key delivery	P7: M -> C	(k)

Table 7.1 provides the definitions of the terms used above.

Field	**Description**
contract	The contract includes a human readable description of the goods being ordered, the price, and the hash of the encrypted goods
TID	Transaction identifier. Globally unique identifier.
L	The log chosen as the point of serialization for this transaction.
t	The token.
expiration	A time after which the transaction, if not completed, is to be aborted.
M	The merchant's identity. Includes account number and merchant bank identifiers as necessary.
k	A symmetric key used only to encrypt the goods.

Table 7.1: Fields in the Anonymous Certified Delivery Protocol

The messages in a transaction using anonymous certified delivery are shown in Figure 7.2.

To initiate the transaction the customer must contact the merchant. Such contact may allow the customer's identity to be inferred as described in Section 2.3.2. However, we assume here that this contact allows the customer to remain anonymous.

During purchase step one, limits on the purchase or special discounts can be negotiated, if the customer wants to include such an offer. For example, the repeated use of a token-specific key would create a pseudonym that would allow for subscriptions or repeat visit discounts with the corresponding inferential disclosure.

At purchase step one, three elements of the transaction must be agreed upon: description of the item, price, and transaction identifier (TID). The transaction identifier must be globally unique. Some combination of a serial number and the merchant's identity can provide a unique TID (as in NetBill).

In preparation for purchase step two the merchant generates a symmetric key, k, and encrypts the requested merchandise with this key. The merchant then sends this encrypted item. Because the merchant encrypts the information, there has been no exchange of valuables. (Note that without a secure channel any observer could steal the merchandise if the transaction is successful.) After the end of this step the token has not been exposed, nor has the customer received a useful item.

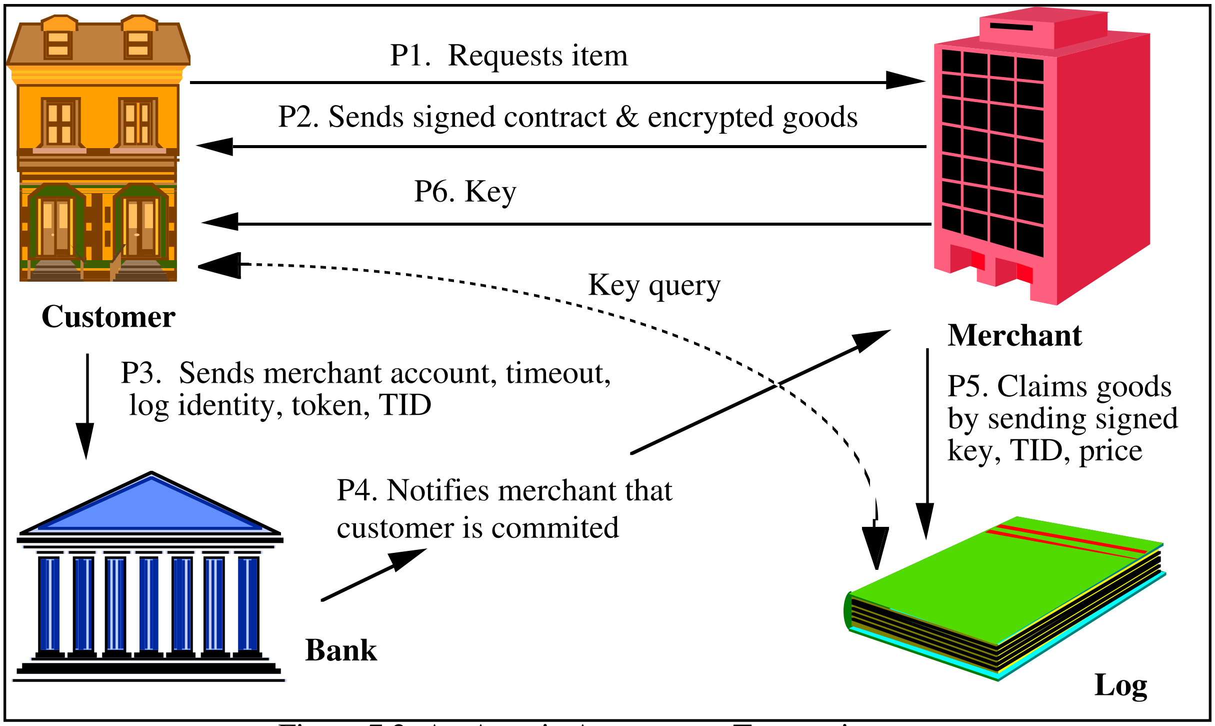

Figure 7.2: An Atomic Anonymous Transaction

In purchase step three the customer commits to the transaction by exposing the token. The customer sends the merchant's bank account number or other appropriate identifier, expiration, the amount of the transaction, the transaction identifier, the log chosen for this transaction, and the token. All this is signed with the token-specific key, q. This signature, together with the receipt ("certificate") from the bank proves that the customer is authorized to spend the token.

In purchase step four the bank notifies the merchant that the customer has committed.

In purchase step five the merchant commits to the transaction by providing the key to the log.

In purchase step six the log generates the necessary receipt to prove the merchant has paid. The customer may demand this receipt in order to obtain the key, if it is not promptly delivered.

In purchase step seven, the merchant completes the transaction by sending the key to the customer. Requiring that the merchant sends the key to the customer reduces the load on the log, since most transactions will be without conflict or failure.

Note that if the transaction is aborted rather than completed the log would generate the following entry and send it to the bank:

P7: L -> B (t, M, TID, failed)l

This message would replace key delivery in the protocol.

If the merchant commits with the log but fails to provide the key, the customer can query the log for the key. The query, shown as a dotted line in Figure 7.2, consists of two messages:

Q1: C -> L	(TID)
Q2: L -> C	$(TID, k)_l$

This protocol provides goods atomicity. If purchase steps one or two, request and contract respectively, fail then there is no transaction.

If purchase steps three through six fail then the merchant is never paid and the customer has only encrypted goods, which are of no use. If purchase steps three or four, customer commitment and notification respectively, fail then the merchant is never paid and the customer never obtains useful goods. If purchase step five, merchant commitment, fails then the key is never delivered, the goods remain encrypted, and the merchant's deposit is never completed. The same is true if purchase step six fails; there is no key delivery and no deposit.

Purchase step six is the global commitment. When purchase step six is completed both key delivery and deposit are assured. Deposit is assured because the bank will credit the merchant in the absence of an abort message. Key delivery is assured because the customer can query the log directly and obtain the key if the merchant fails to deliver it directly. This assures goods atomicity.

Note that there is no need to encrypt purchase message seven: it is simply the key. The key need not be signed for there is already a verifiable copy of the key in the transaction log. The itself key need not be encrypted, since the merchandise was protected by encryption.

The protocol provides certified delivery because the description of merchandise and the item delivered can be verified from purchase step two. This is because that the contract includes a description of the item requested and a checksum of the item delivered, and is accompanied in purchase step two by a unique transaction identifier. All three fields are digitally signed by the merchant. If upon decryption the item delivered does not match the description, then the customer can obtain a refund.

The merchant could send a bogus key and claim the deposit. In this case the holder of the token-specific public key could present the actual merchandise as delivered, and show that the decrypted merchandise does not match the description.

The bank would refund the customer's money by generating a token of equal value and transmitting it to the customer using the anonymous public key for secure delivery. The value of this token would be debited from the merchant’s account.

If the transaction expires, the customer can obtain a new token from the bank.

While there are many places where a dishonest participant or saboteur could delay progress or prevent commitment, there is only one collusion which may result in undetectable fraud. For this reason, there is one trust assumption required by the protocol: the merchant must trust the log to record received messages. If the log, in collusion with the consumer, fails to record the merchant’s commitment, but simply passes the key to the consumer, then the consumer will gain access to the goods while the merchant will not be able to demand payment. This trust requirement is the reason that the merchant selects the log in purchase step two.

The existence of a requirement for trust between the merchant and the transaction log is one argument for combining the log with the bank, since banks already play the role of trusted financial intermediary. In this case, purchase steps two three and four (the contract, customer's commitment and notification, respectively) would not include the log selected, since the selection of the bank would imply a transaction log. Thus, while the merchant would not select the log, the merchant would instead extend trust to the customer's bank.

Conversely, the requirement for trust is also an argument for separating the log and the bank, since this would allow the bank to hold the log to contractual requirements for responsiveness guarantees. Banks could also work as filters with a function analogous to the current role of acquiring banks in the credit card system, and refuse to work with logs which were the subject of many complaints or suspected of fraud.

7.2 Security

There are four public key sets in this system: the transaction log's, the token-specific, the bank's and the merchant's. The value of the bank's key depends upon the technique chosen to generate anonymous tokens.

If an attacker obtains the token-specific private key then the attacker has obtained deposit authorization. However, without the token as well the attacker cannot authorize a transaction.

If an attacker obtains a merchant's key then an attacker can cause the merchant much difficulty in terms of resolving conflicts but cannot force refunds or obtain free merchandise. The attacker can construct a purchase order with a product description, price and faked record of a delivery. (This would require only that the attacker has a token and the corresponding keys. Presumably these are easy to obtain legitimately.) Then the attacker can create a false key, endorse the faked purchase order and send it as a commit. It would appear that the merchant unsuccessfully attempted to defraud the customer. The attacker can then request that the merchant prove the merchandise. The merchant could then detect the attack, decline the transaction, and refuse the transaction amount. There would be no deposit in the merchant's account, so the merchant would owe neither money nor the key.

If an attacker could obtain the keys of the transaction log, the attacker could masquerade as the log. In that case the attacker could steal merchandise as previously described.

7.3 Privacy

The information available to each party in an anonymous certified delivery transaction is shown below in Table 7.2.

Note that law enforcement would only have access to information about the items purchased if their merchant kept transaction-specific data.

In this as in previous cases information during browsing is not included. Recall that to assure anonymity the customer must use an anonymizer or remailer.

Information Party	Merchant	Customer	Date	Amount	Item
Merchant	Full	None	Full	Full	Full
Customer	Full	Full	Full	Full	Full
Transaction Log	Full	None	Full	None	None
Law Enf. w/warrant	Full	None	Full	Full	*Full*
Bank	Full	None	Full	Full	None
Electronic Observer	Full	None	Full	None	None

Table 7.2: Information Available with Anonymous Certified Delivery

7.4 Regulatory Issues

This protocol remove the need for information for the purposes of reliability. However, it does address the needs for information for other purposes, particularly the needs of law enforcement. This protocol has the weaknesses of all anonymous protocols, from the perspective of law enforcement, i.e. there is no information about the customer available to law enforcement. This means that anonymous certified delivery would not be acceptable for legal reasons for transactions over $3,000. There are also constraints on transaction which cross borders. A combination of smurfing and the existence of anonymous bank accounts in foreign jurisdictions can simplify money laundering and tax evasion. Placing limits on transaction size is straightforward; however, limiting international transactions on the Internet is problematic at best.

Issues of reliability and consumer protection are addressed in this protocol. The use of variable public keys limits the amount a customer can lose if a private key is lost. If a key is lost, both the customer and the thief could ‘prove’ ownership of a token. Thus refunds are not possible in the case of lost keys. Customer loss could be limited by requiring that keys be changed at a given increment, or limiting the denomination of a single token if no keys are reused.

The anonymous certified delivery protocol fulfills the requirements for receipts and billing found in the Electronic Funds Transfer Act as implemented in Regulation E. Banks using anonymous certified delivery can provide aggregate information and records on individual withdrawals and deposits.

7.5 Further Questions

The weakness of this protocol is that a public key is needed for each new token. Of course, since the customer is the only one with the private key there is no reason that a single public key could not be used multiple times. In that case the customer's public key resembles a pseudonym. This means that the probability that anonymity is maintained is a function of the frequency of use, duration of use, and breadth of use. That is, the probability of linking any pseudonym, in this case the public key, to real identity decreases as the pseudonym is used in many locations, and as if used frequently or over a long time period. The use of a merchant-specific pseudonym would provide credentials whereby a repeat customer could obtain the appropriate discounts or a subscriber could obtain periodicals.

It may provide an advantage to combine the transaction log and the bank.

Future research issues include probabilistic analysis of the use of pseudonyms, more efficient yet general versions of this protocol and versions of this protocol optimized for different, specific on-line anonymous protocols. One optimization would turn this anonymous certified delivery layer into an anonymous certified delivery protocol by making the public key itself the token (Camp, Harkavy, Tygar and Yee, 1996).

8 Conclusions

“What concerns everyone can only be resolved by everyone.”
Durrenmatt, 1963

This work began with the supposition that a consumer can lose both money and privacy in Internet commerce. Through the consideration of representative commerce systems I have illustrated many of the myriad ways that consumers might lose money and privacy. The weaknesses embodied in each system suggested sometime subtle and sometimes striking problems with the current policy as it affects technical problems with Internet commerce. Now I conclude with a reconsideration and compilation of the problems in Internet commerce and their potential solutions.

This is a brief enumeration of the problems identified in general in the first two chapters and in specific cases in the previous three chapters.

In this work I have removed the technical necessity of trading anonymity and reliability. I have shown that there is a way to provide the highest level of atomicity in an anonymous system. Trading anonymity and reliability is not a legitimate technical option. Whatever limits are put on anonymous transactions are necessitated by social or policy reasons: identity information is not a requirement for reliable transactions or fraud prevention. Any message can fail; any record once created can be copied. Every system should be designed with these facts in mind. Anonymous certified delivery addresses the policy implications of both of these technical realities.

In addition I have brought forward the multiple threats to privacy which converge in Internet commerce. Through an analysis of a set of example protocols I have extracted the implicit policy assumptions of multiple, representative Internet commerce proposals. In order to do so I have developed a formal technique for determining information exchanged in a commercial Internet transaction and shown how such a technique can be useful for examining atomicity.

A further contribution has been the elicitation of system design requirements that simple reporting requirements create, and hopefully therefore the potential for greater understanding among policy makers of the technical implications of their decisions.

Throughout these analyses I have developed a series of policy proposals which I present in the following paragraphs.

Requirements for receipts, billing, contracts and nonrepudiation should reflect the potential embodied in anonymous certified delivery, as well as other technological capabilities. The recognition that consumer identity is not necessary for nonrepudiation should be reflected in record-keeping requirements. For example, requirements created by the Truth in Lending Act and the Electronic Funds Transfer Act as specified in Regulations Z and E, respectively, should be modified to assure that anonymous dispute resolution is supported.

Requirements for dispute resolution information implies a requirement for atomicity. If this is the case in electronic commerce systems, then such a requirement should be explicitly codified; if not, then this too should be made explicit.

In the creation of reporting requirements, I propose the following policy: any data compilation required by law should be protected by law against secondary use. Governmental requirements for information availability should reflect the threat of data surveillance as well as the risks of anonymity. The European model of data commissions provides an excellent example of this approach.

I argue that the previous analyses have shown that the Code of Fair Information Practice should apply to all transactions, and in particular, consumers should know of records as they are being created and be able to opt out. Just as companies offering credit should be required to explain the charges, companies offering network software should be required to explain what information about the user is made available to their software. Consumers should not be required to maintain packet snoopers on their own machines to determine what information is being accessed by the companies and individuals with which they interact. Consumer protection should not depend upon leaked beta software, serendipitous cooperation between hackers and journalists, or the occasional consumer outrage.

Certainly the vast majority of users on the Internet know that there are programs beyond their understanding and protocols foreign to their experience. Users of the Internet know that their information is sometimes transmitted across the globe. Yet there is no way for any but the most technically savvy consumer to determine just what information she leaves behind as she travels the information infrastructure. There is a definite need for consumer privacy protection on the Internet.

Potentially the most contentious conclusions in this dissertation are that protections for privacy need to be expanded, and the constraints on the export of encryption technologies should be decreased. In particular I have argued that controls on secondary disclosure should be extended. The Right to Financial Privacy Act should be strengthened, and should be extended to apply to all institutions that compile consumer financial data.

These conclusions will be viewed in very disparate ways by different communities. Thus here I consider three possible perspectives in the following section: the perspectives of law enforcement, the business community, and the civil libertarian.

8.1 Law Enforcement

Government has a need for information in order to fulfill its legitimate purposes. Perhaps the greatest source of conflict between privacy and data availability has been in law enforcement. Anonymity is generally opposed by law enforcement. The law enforcement community is charged with assuring that identifiable individuals are held responsible for specific acts. Detecting patterns of illegal activity and pursuing the appropriate parties is made easier by increasing data availability -- thus the range of reporting requirements created for the needs of law enforcement.

The opposition to anonymity by law enforcement has not been absolute. For example, Mark Twain Bank has been offering anonymous Digicash to consumers. The approval of Digicash for use in the United States is based on two factors: there exist size limits on anonymous transfers and Digicash can only be used once before deposit. This means that Digicash can go through a single transaction, but cannot go through a chain of transactions. Thus in every transaction Digicash returns to fully traceable banking channels.

The law enforcement community has a set of data requirements that are made explicit in carefully crafted requirements for data as explained in Chapter 2. These have included limits to scale in anonymous transactions . However, the requirements of law enforcement have not prevented the general use of anonymous electronic currency.

The use of anonymity and encrypted communication can allow widely distributed individuals to plan and implement illegal activities without any fear of surveillance. Many in the pro-crypto, libertarian community in fact advocate the removal of international borders. However, even making international borders porous has resulted in an inability to contain regional conflicts, and separatist conflicts may result in the deaths on another continent. Internationally interconnected networks have magnified the ability of one individual to cause harm (Baird, 1996).

Like legitimate businesses, criminal organizations are becoming leaned and meaner. Criminal organizations are capable of advanced administration. Because of computing and communications technologies criminal organizations need fewer people, and are therefore more difficult to penetrate (Bickford, 1996). Thus the ability of these organizations to move money without a trace makes observing their actions, or even locating them, difficult.

International criminal organizations may be assisted by criminal governments, thus Americans cannot always depend on foreign governments to protect their interests. Law enforcement and national security are increasingly interdependent, and the lack of coordination and information can be costly. Collapsing empires result in the rise of organized crime to enforce property and contract rights that cannot be enforced by the government, and these organizations can create international corruption (Rodman, 1996). Currently there are over 250 Russian criminal organizations with international organizations (Bickford, 1996).

Basic safeguards such as a prohibition on anonymous bank accounts and limited anonymity in purchases fulfill the needs of law enforcement. Anonymous electronic funds transfer mechanisms should be evaluated with the reality of money laundering in mind. Five hundred billion dollars in laundered funds comes from the United States, with eighty percent of that being drug money (Bickford, 1996). The ease of smurfing makes the traditional simple limits on funds transfers inadequate in the electronic realm (Office of Technology Assessment, 1995). The protocol presented in the previous chapter resolved conflicts between anonymity and the need for information for reliability; but it did not resolve conflicts between anonymity and the need for information for law enforcement purposes.

8.2 The Business Community

There exists a profitable market for secondary uses of consumer information. Companies profit from both internal use and the ability to market data. Both companies and consumers profit when targeted advertising results in a transaction. Consumer data is most often used by businesses to better serve the needs of customers and to identify financial opportunities of which the customer may be unaware. The protection of privacy must be more valuable than the uses of consumer data for privacy protection to be the choice of the business community. Yet once that case is made, the market will move to protect privacy.

Regulatory limits on the use of consumer transactional data would create an economic loss to marketers. Thus many merchants, including those that sell financial data, would oppose regulatory limits on the collection, analysis and disclosure of consumer information. Because consumer information is necessary in a modern economy there has been no proposal to prohibit all consumer data flows. Limiting the flow of consumer data could have unforeseen effects on an information economy.

Market resistance to consumer privacy protections has not been uniform. For example, Microsoft has pledged to follow the standards set forth in the European privacy directive as described in Chapter 3. Netscape addressed the possible misuse of anonymous ftp as quickly as possible in the release of Netscape Navigator 2.01. A possible increase in the use and trust of network services could profit such companies if the consumer could be assured that any company will not surreptitiously obtain additional data during transactions. No network services provider or business profits when its customers are subject to third party surveillance. Thus companies may cooperate to create a floor of minimum privacy protection, and provide the equivalent of the Better Business Bureau seal to assist their customers in choosing trustworthy merchants.

The business community has and will provide anonymity to those willing to purchase such a service. Those willing to invest time and effort in such a search can find credit cards with varying policies for secondary disclosure of consumer transactional information. On the Internet, those interested in anonymity can use Digicash, for the price of the cost of Mark Twain's services and possible merchant fraud.

It is possible that externalities exist with widespread implementation of anonymity. That is, there may exist a critical mass in terms of number of consumers who adopt anonymity before there are increasing returns to scale in distributing anonymous software or providing anonymity through secure intermediaries. Thus it is possible that new price paradigms which recognize the existence of positive network eternalities are needed. If this is the case, and there is a powerful market for privacy-protecting services, it may yet be served by market forces. Many customers may be willing to pay a price, but not the premium that Mark Twain would charge, for privacy.

Opposition from the business community to anonymity is economic and not ideological. There are long term benefits which are possible drivers of adoption of anonymity such as increased use of financial services and improved customer relations. The profit motive can serve to provide privacy-enhancing services for consumers willing and able to purchase these services.

The business community has a fairly uniform perspective on the prohibition of the export of cryptography (United States Council for International Business, 1993). Export of encryption allows producers of software, hardware and systems to take advantage of a traditional American strength in serving the global market. The prohibition of export of strong cryptography hurts business by preventing them from effectively serving these markets. Thus the business community supports the free export of cryptography, which enables anonymity.

Anonymous certified delivery as described here provides all the information necessary for primary business purposes but does not provide information for secondary uses.

8.3 Civil Libertarians

Civil libertarians are both strong advocates of privacy, and strong supporters of unrelated social goals.

Civil libertarians are concerned about both the potential for surveillance and the effect of a perception of surveillance. Consider the impact of the proposal put forth by law enforcement for assuring access to clear text of all communications using key escrow. The key escrow proposals effectively assure government access to information which will increasingly be financial transactions. The establishment of a governmental EFT service was considered and rejected by the Privacy Commission. The Commission's objection to the creation of a government-operated EFT system was that such a system would result in government surveillance, and thus enable government to easily prescribe 'correct' behavior (Privacy Protection Commission, 1977). Key escrow for access to consumer financial transactions poses the same threat.

Secondary disclosure of information includes disclosure to the government. In Lamont v Postmaster General the Supreme Court noted that observation by the Federal Government has a chilling effect on the pursuit of information. There is no reason that this will change as information becomes electronic and not paper-based. Lamont v Postmaster applied both to free and purchased information. Civil libertarians are fighting for a recognition that this principle in law applies to electronic transactions.

From the civil libertarian perspective, law enforcement requirements have only served to limit the availability of security and privacy through constraints on cryptography.

The prohibition of exporting encryption technology has had ubiquitous effects. This prohibition effectively prevents strong cryptography for the protection of privacy, as discussed in anonymous certified delivery, from being implemented. The use of public key Kerberos would be prohibited. The advantage of public key Kerberos is that no central key authority is required -- and it is precisely the lack of a central key authority that causes law enforcement concerns.

Civil libertarians note that with the proliferation of information technology, cryptography is no longer a predominantly military technology. The list of uses for cryptography more resembles the broad range of applications for internal combustion than the narrow focus of ballistic missile technology. Cryptography is used in every electronic commerce system. Escrow prohibition has prevented security from being an integral part of operating systems and software for Internet access from desktop machines, and thus limited privacy.

From the perspective of civil libertarians, that law enforcement is constrained from unreasonable search and seizure does not mean that citizens have to live with a network designed to make reasonable search and seizure simple. Citizens still have the right to avoid law enforcement access, without a presumption of guilt.

Civil libertarians are concerned about use of consumer data by the business community as well as government. The lack of information for secondary use is a strength and not a weakness from this perspective. Civil libertarians would applaud constraints on secondary use of consumer data. However, civil libertarians also recognize the need for data to meet social needs, such as preventing discrimination, and thus support some federal oversight. Civil libertarians also seek to protect consumers' economic rights, so concerns about reliability will affect their support for privacy. Civil libertarians are the most likely supporters of advanced but expensive technical solutions, such as anonymous certified delivery, to problems of privacy and reliability.

All civil libertarians support the removal of constraints on the use of strong cryptography even for international discussions. While the fight for consumer privacy may often result in conflict between civil libertarians and the business community, they are united in their opposition to the prohibition of the export of strong cryptography.

8.4 In Closing

Before this work, and unless it is implemented, consumers had to choose between allowing others to know the contents of their wallets and risking that others would steal those contents. This is no longer true. Now limits on privacy are a result of governmental requirements for information.

Statutes dealing with fraud have addressed the threats the consumer faces to her wallet; threats the consumer faces to her privacy also need to be addressed. Consumers transmit xinformation without consent, or even knowledge. The risks of information disclosure are increasing. Browsing the Web can offer a heady feeling of anonymity, of walking without footprints across information space. In truth, information about users of Internet services is easy to gather.

The shadow of Olmstead v United States (Olmstead v United States, 1928), which stated that early users of telephone technology had no expectation or right to privacy, hangs over the Internet. Hopefully we will not suffer the 39 years of data surveillance allowed by slow adaptation of the law to new technologies, as suggested by the telephony model.

Abuses of Internet user information extend beyond the identity and transactional information addressed here. Agents search the Internet for every available email address, compiling marketing lists which one has no right to access. Searchable databases of Usenet posts in various groups assure that the individual has extremely limited rights on the dissemination of “his religion, ideology, opinions, and interests” (Douglas, 1974). That surveillance is enabled by the technology does not mean that it is an inevitable effect of technology. Policy can prevent the surveillance potential of technology from becoming a reality; at the least it can empower consumers to have knowledge of the information they are disclosing. For these policies to be adopted consumers must also act as citizens.

In the long run it is uncertain whether the privacy traditions of libraries or the disclosure traditions of banks will dominate when financial data is linked to reading habits and information access. There are fundamentally different legal and professional traditions in the information provision and financial services sectors. Regardless of the dominant tradition, the sheer distance between these traditions suggests that it is important to build in the capacity to change, rather than cementing de facto standards.

According to current law, under United States v Miller, financial information belongs to the service provider, and the consumer has no privacy interest in her records. It is my hope that, as electronic commerce becomes ubiquitous, United States v Miller will become as United States v Olmstead: quoted for the passion of the dissent and regarded as an instructive case of error by the courts. Yet unless and until five of the nine Justices agree, it is the shared responsibility of policy makers, bankers, brokers, computer scientists and consumers to assure that America is not hard-wired with a surveillance infrastructure.

9 Bibliography

5 USC §552 Privacy Act
12 USC §1829 Bank Secrecy Act
12 USC §1829 Money Laundering Act
12 USC §2903 Community Reinvestment Act
12 USC §3403 Financial Privacy Act
12 CFR §202 Home Mortgage Act
15 USC §1601 Truth In Lending Act
15 USC §1691 Equal Credit Opportunity Act
15 USC §1692 Fair Debt Collection Practices Act
15 USC §1694 Electronic Funds Transfer Act
18 USC §1029 Computer Fraud and Abuse Act
22 USC §2571 Arms Control Act
22 CFR §121 International Traffic in Arms Regulation
26 USC §6103, 31 USC §3711 Debt Collection Act
31 CFR §103 Know Your Customer Requirements
35 USC §3401 Right to Financial Privacy Act
42 USCS §3608, 15 USC §1681, 12 USCS §1708 Fair Credit Reporting Act
49 USC §1666 Fair Credit Billing Act
50 USC 2401 Export Administration Act

Anderson, R. E., Johnson, D.G., Gotterbarn, D. and Perrolle, J., 1993, "Using the ACM Code of Ethics in decision making," *Communications of the ACM*, Vol. 36, 98-107.

Alderman, E. and Kennedy, C., 1995, *The Right to Privacy*, Alfred A Knopf, New York, NY.

Axsmith C., 1992, "Email privacy and the law," *Proceedings of the 15th National Computer Security Conference*, Baltimore MA, 120-125.

Baird, Z., 1996, "How have other nations balanced legal an national security threats and responded to a changed world?" *American Bar Association Standing Committee on Law and National Security Law Enforcement and Intelligence Conference*, September 19.

Baker, 1994, *A Concise Introduction to the Theory of Numbers*, Cambridge University Press, New York, NY.

Bernam J., 1991, "Establishing a legal framework for freedom and privacy on the electronic frontier," *Conference on Computers Freedom and Privacy*, Washington D.C.

Bickford, D., 1996, "The changed threat to US national security -- new problems and priorities," American *Bar Association Standing Committee on Law and National Security Law Enforcement and Intelligence Conference*, September 19.

Bloustein, 1968, "Privacy as an aspect of human dignity: an answer to Dean Prosser," *New York University Law Review*, Vol. 39, 962-970.

Brands, S., 1993, "Untraceable off-line cash in wallet with observers," *Advances in Cryptology - CRYPTO '93,* Springer-Verlag; Berlin, 302-318.

Brennan, 1989, *Florida v Riley*, 488 U.S. 445, 466 (Brennan J., dissenting).

Brickell, E., Gemmell, P., and Kravitz D., 1995, "Trustee-based tracing extensions to anonymous cash and the making of anonymous change," *Proceedings of the Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, San Francisco, California, 22-24 January, 457-466.

Britt, P., 1994, "Moving forward with smart cards," *Savings & Community Banker,* Vol. 3, No. 11, 6-7, November.

Business Week, 1993, "ATM shouldn't stand for 'artfully taken money,'" *Business Week* (Industrial/Technology Edition), May 31, 110.
Bureau of Census, 1995, *Statistical Abstracts of the United States 1995 (115th Edition)*, Department of Commerce, Washington, D.C.
Camp, L.J. and Tygar, J.D., 1994, "Providing auditing while protecting privacy," *The Information Society*, Vol. 10, 59-71, March.
Camp, L. J. ,Sirbu M. and Tygar, J. D., 1995, "Token and notational money in electronic commerce," *Usenix Workshop on Electronic Commerce*, July, New York, NY.
Camp, L.J., Harkavy, M., Tygar, J.D. and Yee, B., 1996, "Anonymous atomic transactions," *2nd Annual Usenix Workshop on Electronic Commerce*, November, Oakland, CA.
Cerf, V., 1993, "How the Internet came to be," *The On-line User's Encyclopedia*, ed. B. Aboba, Addison-Wesley, New York, NY.
Clark G. and Acey M., 1995, "Mondex blows users anonymity," *Network Week* (U. K.), Vol. 1, No. 8, Col. 1, October 25.
Chaum, D., 1985, "Security without identification: transaction systems to make big brother obsolete," *Communications of the ACM*, Vol. 28, 1030-1044, October.
Chaum, D., 1989, "On-line cash checks," *Advances in Cryptology - EUROCRYPT '89*, 288-293.
Chaum, D., 1992, "Achieving electronic privacy," *Scientific American*, Vol. 267, 76-81.
Chaum, 1994, *Prepaid Smart Card Techniques: A Brief Introduction and Comparison*, Digicash, Holland.
Chaves, C., 1992, "The death of personal privacy," *Computerworld*, 25 - 27, January.
CommerceNet, 1995, *The CommerceNet/Nielsen Internet demographics survey: Executive Summary*, CommerceNet, http://www.commerce.net/information/surveys/toc.html, October 30.
Compaine B. J., 1988, *Issues in New Information Technology*, Ablex Publishing; Norwood, NJ.
Computer Science and Telecommunications Board, 1994, *Rights and Responsibilities of Participants in Networked Communities,* National Academy Press, Washington, D.C.
Cox, B., 1994, *Maintaining Privacy in Electronic Transactions,* Information Networking Institute, Carnegie Mellon University, Pittsburgh, PA.
Cox, B., Tygar, J.D. and Sirbu, M., 1995, "NetBill security and transaction protocol," Usenix *Workshop on Electronic Commerce*, July, New York, NY.
Cross Industry Working Group, 1995, "Electronic cash, tokens and payments in the National Information Infrastructure", *http://www.cnri.reston.va.us:3000/XIWT/documents/dig_cash_doc/ToC.html,* September.
Davies, 1981, *The Security of Data in Networks*, IEEE Computer Society Press: Los Angeles, CA.
Davis, P., 1995, "Senate Republicans say the Earned Income Tax Credit is becoming too expensive," *National Public Radio Morning Edition,* National Public Radio; number quoted by Margaret Richardson, Commission of the Internal Revenue Service, August 17 (also available on http://www.realaudio.com/contentp/npr/nb0817.html).
Denning, D., 1982, *Cryptography and Data Security*, Addison-Wesley Publishing; Reading, MA.
Department of Defense, 1985, *Department of Defense Trusted Computer System Evaluation Criteria*, National Computer Security Center, Fort George G. Meade, MD.

Diffie W. and Hellman M. E., 1976, "New Directions in Cryptography," *IEEE Transactions on Information Theory*, Vol. 7, November, 644-654.
Diffie, W. and Hellman, M. E., 1979, "Privacy and authentication: an introduction to cryptography," *Proceedings of the IEEE*, Vol. 67, 18-48.
Douglas, 1974, *California Bankers Association v Schultz,* 416 U.S. 21,85, 94 S. Circuit, 1494, 1529, 39 L. Ed. 2d 812, dissent.
Draper, S., 1989, "Security aspects of smart cards," *Computer Security in the Age of Information*, Elsevier Science Publishers B.V., Holland, ed. Caelli.
Duncan G. and Lambert D., 1986, "Disclosure-limited data dissemination," *Journal of the American Statistics Association*, Vol. 81, 10-27.
Duncan G. and Lambert D., 1989, "The risk of disclosure for microdata," *Journal of Business and Economic Statistics*, Vol. 7, 207-217.
The Economist, 1996, "Who's who on the Internet," *The Economist*, Vol. 340, No. 7976, July 27.
Echikson, W., 1994, "French risk it all on a smart card," *Boston Globe*, February. 28, 17:2.
FAIR, 1996, *FAIR Media Bias Detector*, http://www.igc.apc.org/fair/media-bias-detector.html, April 15.
Federal Bureau of Standards, 1977, *Federal Information Processing Standards Publication 46: Announcing the Data Encryption Standard*, US Government Printing Office: Washington, DC.
Federal Communications Commission, 1995, *Telephone Subscribership in the United States,* US Government Printing Office: Washington, DC, April.
Federal Reserve Bank of New York, 1996, "Regulation E - Electronic Funds Transfer - revisions to regulation and official staff commentary," *Federal Register*, Vol. 61, No. 86, May 2.
Feige, U., Fiat, A. and Shamir, A., 1987, "Zero knowledge proofs of identity," *Proceedings of the 19th ACM Symposium on Theory of Computing*, 210-217.
Fenner, E, 1993, "How mortgage lenders can peek into your files," *Money,* 44-48, April.
Financial Service Technology Consortium, 1995, *Electronic Payments Infrastructure: Design Considerations*, http://www.llnl.gov/fstc/projects/commerce/public/epaydes.htm, November.
First Virtual, 1995a, *Information About First Virtual,* http://www.fv.com/info, October 8.
First Virtual, 1995b, *The Fine Print,* http://www.fv.com/info/terms.html, June 24.
Fischer, M.J., 1988, "Focus on industry," *Journal of Accountancy*, 130-134, June.
Freier, A., Karlton, P. and Kocher, P.C., 1996, *The SSL Protocol, Version 3,* Netscape Communications Corporation, Mountain View CA. Also available at ftp://ietf.cnri.reston.va.us/internet-drafts/draft-freier-ssl-version3-01.txt.
Froomkin, A.M., 1995, "Anonymity and its enmities," *Journal On-line Law*, Vol. 1, No. 1, art. 4.
Froomkin, A. M., 1996, "Addressing law enforcement concerns in a constitutional framework," *SAFE: Security And Freedom through Encryption Forum*; Palo Alto, CA. July 1, 1996.
Fuller, T., 1732, *Gnomologia,* London, U.K.
Garfinkle, S. and Spafford, G., 1986, *Practical UNIX Security, Second Edition,* O'Reily & Associates, Sebastopol, CA.
Goradia, V., Kang, P., Lowe, D., Magruder, P., McNeil, D., Mowry, B., Panjwani, M., Somogyi, A., Wagner, T. and Yang, C., 1994, *NetBill: 1994 Prototype*, Carnegie Mellon University; Pittsburgh, PA. Available as INI technical report INI TR 1994-11.
Gray, J. and Reuter, A., 1993, *Transaction Processing: Concepts and Techniques*, Morgan Kaufmann Publishers; San Francisco, CA.

Halpern, 1991; “Rethinking the right of privacy: dignity, decency and the law’s limitations,” *Rutgers Law Review,* Vol. 43, No. 3, 539-563.
Hansell S., 1995, “Mastercard joins banks to plan card that works like cash,” *The New York Times*, August 17, 1995, D2.
Hanushevsky, A., 1995, *Electronic Commerce Pag*e, http://abh.cit.cornell.edu/ecom.html, November.
Hart, A.S., 1996, personal communication via email, VP Software Engineering, CyberCash, Inc., Reston, VA, May 16.
Harrison, C., 1994, “Shoppers urged to guard against credit card fraud,” *Atlanta Constitution*, December 27, C4:5.
Harvard Law Review, 1991 “Addressing the new hazards of the high technology workplace,” *Harvard Law Review,* Vol. 104, 1898-1916.
Heggestad, A., 1981, *Regulation of Consumer Financial Services*, Abt Books, Cambridge, MA.
Henry v Forbes, 1976, 433 F. Supp. 5.
Hodges, A., *Alan Turing: The Enigma*, Simon & Schuster, Great Britain, 1983.
Hoffman, L, Kalsbeek, W.D. and Novak, T.P., 1996, *Internet Use in the United States: 1995 Baseline Estimates and Preliminary Market Segments*, Project 2000 Working Paper, April 12. Also available at http://www2000.ogsm.vanderbilt.edu/baseline/1995.Internet.estimates.html.
Hoffman, L. and Clark P., 1991, “Imminent policy considerations in the design and management of national and international computer networks,” *IEEE Communications Magazine*, February, 68-74.
Ingramham, D. G., 1991, “Coming of age in cyberspace,” *Conference on Computers Freedom and Privacy*, Washington DC.
Internet Domain Survey (IDS), 1995a, *Exponential Growth of the Number of Computers Connected to the Internet*, http://www.nw.com/zone/WWW/top.html, November.
Internet Domain Survey (IDS), 1995b, *Hosts Stats by County,* http://www.nw.com/zone/WWW/isoc-pr-9501.txt, November.
Jennifer, G., Steiner, B., Neuman C., and Schiller.J.I., 1988, “Kerberos: an authentication service for open network systems,” *Proceedings of the USENIX Winter Conference,* February, 191-202.
Johnson, B.S., 1989, “A more co-operative clerk: the confidentiality of library records,” *Law Library Journal*, Vol. 81, 769-804.
Johnson, D., 1989, “Documents disclose F.B.I. investigations of some librarians,” *New York Times,* November 7, A, 1:1.
Johnson, K., 1993, “One less thing to believe in: fraud at fake cash machine”, *New York Times*, May 13, A, 1:5.
Kailer, 1995, “Reasoning about accountability in protocols for electronic commerce,” Proceedings of the IEEE Symposium on Security and Privacy, Oakland CA, May 1995.
Kalven, 1966, “Privacy in tort law - were Warren and Brandeis wrong?” *Law & Contemporary Problems*, Vol. 31, 326-332.
Kaplan, E.H., 1991, “Needles that kill: Modeling human immunodeficiency virus transmission via shared drug injection equipment in shooting galleries,” *Reviews of Infectious Diseases*, Vol. 11, 289-298.
Karasik E.,1990, “A normative analysis of disclosure, privacy and computers: the state cases,” *Computer Law Journal*, Vol. 10, 603-634.
Katz v United States, 1967, 389 US 351, 369 F2d 130 (9th Cir).
Kaylin J., 1992, “When the needles do the talking,” *Yale,* April, 34-37.
Lamont v Postmaster General, 1965, 381 U.S. 301, 301.
LaPlante, A., 1994, “Citibank's smart move,” *Information Week*, Vol. 492; No. 12 September, 42.

Lewis, T., 1996, personal communication, T. Lewis is Project Manager of Technology Research at Visa.
Low, S., Maxemchuk, N.F. and Paul, S., 1993, "Anonymous credit cards," *First ACM Conference on Computer and Communications Security,* November.
Madsen, W., 1992, *Handbook of Personal Data Protection*, Stockton Press; New York, N.Y.
Markoff, J., 1995, "Security flaw is discovered in software used in shopping," *The New York Time*s, September 19, A1, D21.
Mastercard, 1995, *Secure Electronic Payment Protocol Specification Draft Version 1.1*, http://www.mastercard.com/Sepp/sepptoc.htm, November, Part 2.
Mastercard, 1996, *Secure Electronic Transaction Technology*, *Draft*., http://www.mastercard.com/SETT.
Mayland, P. F., 1993, "EFT network risk begs CEO attention," *Bank Management*, No. 69, Vol. 10, 42-46, October.
McClellan, D.,1995, "Desktop counterfeiting," Technology *Review*, http://web.mit.edu/afs/athena/org/techreview/www/articles/feb95/mcclellan.html, February/March.
McGraw D., 1992, "Facing the specter of AIDS" *Boston Globe*, March 13, 3-5.
Medvinski, G. and Neuman, B.C., 1993, "NetCash: A design for practical electronic currency on the Internet," *Proceedings of the First ACM Conference on Computing and Communications Security*, November.
Miller, B.C., Neuman, C., Schiller, J.I. and Saltzer, J.H., 1987, "Section E.2.1: Kerberos authentication and authorization system", *MIT Project Athena*, Massachusetts Institute of Technology, Cambridge, MA., December.
Miller M.W., 1992, "Data tap: patients' records are treasure trove for budding industry," *Wall Street Journal,* February 27, A 1:6.
Morgan G., 1992, "Balancing national interest," *The Institute,* Vol. 16, November/December.
Mosteller, F. 1965, *Fifty Challenging Problems in Probability with Solutions*, General Publishing Company, Ltd.: Toronto, Canada.
Mundt K. H., 1992, "New dimensions in data security," *Proceedings of the 15th National Computer Security Conference*, Baltimore MA, 438-447.
NAACP (National Association for the Advancement of Colored People) v Alabama, 1958, 357 US. 449.
National Bureau of Standards, 1977, *Federal Information Processing Publication 46: Specifications for the Digital Encryption Standard,* United States Government Printing Office; Gaithersburg, MA.
National Center for Supercomputing Applications, 1995, *NCSA Mosaic Web Index,* http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/Docs/web-index.html, November.
National Computer Security Center, 1985, *Trusted Systems Evaluation Criteria DOD-5200.28-STD*, United States Government Printing Office; Gaithersburg, MA.
National Computer Security Center, 1990, *Trusted Network Interpretation Environments Guideline NCSC-TG-011*, United States Government Printing Office; Gaithersburg, MA.
National Institute of Standards and Technology, 1991, *Proposed Federal Information Processing Standard for Digital Signatures,* Federal Register, Vol. 56, August, 42980-42982.
National Institute of Standards and Technology, 1994, *Federal Information Processing Standards Publications 185: Escrowed Encryption Standard*, United States Government Printing Office; Gaithersburg, MA.
National Research Council, 1996, *Cryptography's Role in Securing the Information Society*, National Academy Press, Washington, DC.

Netscape, 1996, *Netscape Commerce Server*, http://home.netscape.com/comprod/netscape_commerce.html, May.
New York Times, 1995a, "Woman missing bank card finds she is overdrawn $346,770", *New York Times*, Feb. 12, 1, 36:1.
New York Times, 1995b, "Credit union's error is thieves' delight," *New York Times*, Feb. 9, B, 9:6.
Newberg, P., 1989, *New Directions In Telecommunications Policy*, Duke University Press; Durham, NC.
Nimmer, R. T., 1992, *The Law of Computer Technology,* Warren, Gorham & Lamont, Boston, MA.
Okamoto, T. and Ohta, K., 1991, "Universal electronic cash," *Advances in Cryptology-CRYPTO '91*, Springer-Verlag; Berlin, 324-336.
O'Keefe, M., 1994, "Portable POS debit terminals mean greater convenience," *Bank Systems & Technology*, Vol. 31, No. 11, 35-37, November.
Office of Technology Assessment, 1985, *Electronic Surveillance and Civil Liberties OTA-CIT-293*, United States Government Printing Office; Gaithersburg, MA.
Office of Technology Assessment, 1986, *Management, Security and Congressional Oversight OTA-CIT-297,* United States Government Printing Office; Gaithersburg, MA.
Office of Technology Assessment, 1995, *Information Technologies for Control of Money Laundering*, OTA-ITC-630, United States Government Printing Office; Gaithersburg, MA.
Olmstead v United States, 1928, 277 US 438, 48 SCt 564, 72 LEd2d 944.
Pool, I., 1983, *Technologies of Freedom,* Harvard University Press, Cambridge MA.
Privacy Protection Study Commission, 1977, *Personal Privacy in an Information Society,* Government Printing Office, Washington, D.C.
Prosser W.L., 1941, *Handbook of the Law of Torts,* West Publishing Co., St. Paul, MN.
Rabin, M. O., 1978, Digital Signatures, *Foundations of Secure Communication*, Academic Press: New York, NY, 155-168.
Reid, M. A. and Madam, M. S., 1989, "IC card design: technology issues," *Information Age,* Vol. 11, No 4, 211-216.
Rivest, R. L., Shamir, A. and Adleman, L., 1978, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, Vol. 21, 158-164.
Rivest, R.L. and Shamir, A., 1996, "PayWord and MicroMint: Two simple micropayment schemes," submitted to Eurocrypt '96.
Rodman, P., 1996, "Loss of national sovereignty and control by nation states," *American Bar Association Standing Committee on Law and National Security Law Enforcement and Intelligence Conference*, September 19.
Rubin, L. and Cooter, R., 1994, *The Payment System: Cases Materials and Issues*, West Publishing Co.; St. Paul, MN.
Sandberg, J., 1995, "Netscape software for cruising Internet is found to have another security flaw," *The Wall Street Journal,* September 25, B12.
Schlossberg, H., 1993, "Victims tired of researchers getting away with murder," *Marketing News*, August 16, A16:1.
Schnorr, C. P., 1990, "Efficient signature generation of smart cards," *Advances in Cryptology-CRYPTO* '89, Spring-Verlag: Berlin, 239-252.
Schneier, B., 1995, *Applied Cryptography*, *Second Edition*, John Wiley & Sons, Inc., New York, NY.
Shamir, A., 1979, "How to share a secret," *Communications of the ACM,* Vol. 22, 612-613.
Simpson, 1996, *The Effects of Electronic Credentials Lifetime on the Risks and Costs of Electronic Commerce*, Qualifier Report, Department of Engineering & Public Policy, Carnegie Mellon University, Pittsburgh, PA.

Sirbu, M., and Tygar, J. D., 1995, "NetBill: an Internet commerce system optimized for network delivered services," *IEEE ComCon,* San Francisco, CA. March 6
Smith, S., 1992, *A Theory of Distributed Time*, Ph.D. dissertation, Carnegie Mellon University. Available as CMU technical report CMU-CS-92-231.
Speiser, S. M., Krause, C.F. and Gans, A. W., 1991, *The American Law of Torts*, Clark Boardman Callaghan, New York, NY.
Trublow, G. et al, 1991, *Privacy Law and Practice*, Times Mirror Books, New York, NY.
Trublow, G. 1992, "When Is Monitoring E-Mail Really Snooping?" *IEEE Software*, Vol. 9, No. 2. 97-98, March.
Tunstall, J., 1989, "Electronic currency," *Smart Card 2000: The Future of IC Cards: Proceedings of the IFIP,* Elsevier Science Publishers B.V., Holland, eds. Cahum & Schaumuller-Bichl .
Turn, R. and Ware, W., 1976, "Privacy and security in information systems," *IEEE Transactions on Computers*, C-25, 1353-1361.
Tygar, J. D. and Yee, B., 1991, "Strongbox: a system for self securing programs", *CMU Computer Science: A 25th Anniversary Commemorative*, ed. R. Rashid, 1991, Addison-Wesley and ACM Press; New York, NY, 163-198.
Tygar, J. D., 1996, "Atomicity and electronic commerce," *Proceedings of 1996 Symposium of Principles of Distributed Computing*, ACM Press, Philadelphia, PA.
United Nations, 1995, *The United Nations and Human Rights 1945-1995: The United Nations Blue Book Series. Vol. VII*, United Nations; New York, New York.
United States Council for International Business, 1993, Statement of the United States Council for International Business on the Key Escrow Chip, United States Council for International Business, NY, NY.
United States District Court, 1992, United States v Julio Fernandez, John Lee, Mark Abene, Elias Ladopoulos, and Paul Stira, Indictment 92 CR S63.
United States v Miller, 1976, 425 U.S. 435.
United States v Payner, 1980, 447 U.S. 727, 100 S. Ct. 2439, 65 L. Ed. 2d 468.
Van Natta, D., 1995, "5 phone marketers arrested in credit card sting," *New York Times*, August 15, A, 14:2.
Verisign, 1995, *Verisign Expands Digital ID Offerings To Leading Web Servers*, http://www.verisign.com/pr/pr_servers.html, November
Verisign, 1996, *Frequently Asked Questions About Digital ID's*, http://digitalid.verisign.com/id_faqs.htm, May 26.
Visa, 1995, *Secure Transaction Technology Specifications Version 1.1*, http://www.visa.com/visa-stt/index.html, November.
Wacker, J.,1995, "Drafting agreements for secure electronic commerce" *Proceedings of the WorldWide Electronic Commerce: Law, Policy, Security & Controls Conference*, October 18-20, Washington, DC, 6.
Walden, I., 1995, "Are privacy requirements inhibiting electronic commerce," *Proceedings of the WorldWide Electronic Commerce: Law, Policy, Security & Controls Conference*, October 18-20, Washington, DC, 10.
Warren S. and Brandeis L., 1890, "The right to privacy," *Harvard Law Review*, Vol. 4, 193-220.
Waters v Fleetwood, 1956, 91 SE2d 344, Georgia.
Wood, J.C. and Smith, D.S., 1991, "Electronic transfer of government benefits", *Federal Reserve Bulletin,* Vol. 77, No. 4, 203-217, April.
Woodyard C., 1991, "Lungren joins suit accusing TRW of illegal practices," *Los Angeles Times*, July 9, 1:5.
Yee, B., 1994, *Using Secure Co-processors*, Ph.D. dissertation, Carnegie Mellon University. Available as CMU technical report CMU-CS-94-149.
Ziegler, R. F., Brodsky, D. E. and Sanchez, C. M., 1993, "US securities crime," *International Corporate Law, Criminal Investigations Supplement*, 69-74, May.

Zimmerman, P., 1995, *The Official PGP User's Guide*, MIT Press, Cambridge, MA.
Zuckerman, G., 1994, "Insider trading is back," *Investment Dealers Digest*, Vol. 60, No. 2, 12-15, May 30.