Language Modeling and Encryption on Packet Switched Networks[*]

Kevin S. McCurley

Google

Abstract. The holy grail of a mathematical model of secure encryption is to devise a model that is both faithful in its description of the real world, and yet admits a construction for an encryption system that fulfills a meaningful defintion of security against a realistic adversary. While enormous progress has been made during the last 60 years toward this goal, existing models of security still overlook features that are closely related to the fundamental nature of communication. As a result there is substantial doubt in this author's mind as to whether there is any reasonable definition of "secure encryption" on the Internet.

1 Introduction

In any area of science there is a fundamental tension between the desire to describe the real world with a model that is accurate in detail, vs. the desire to use models that facilitate precise mathematical reasoning. In the case of cryptology, if a model fails to describe the real-world application, it leaves room for attacks in the real world that were not anticipated by the model itself. This has been highlighted in recent years by the discovery of multiple "side-channel attacks" that are very effective against real-world systems, but usually fall outside the scope of existing security models. Examples include algorithmic timing analysis [12], differential power analysis [13], protocol fault analysis [4], and differential fault analysis[3]. There is at least anecdotal evidence that many other side channel attacks exist (e.g., RF and acoustic attacks) for popular models of security. Unfortunately, a security model is successful only to the extent that it accurately describes the process and the capabilities of the adversaries; any omission, oversight, or ambiguity in this may properly be regarded as a weakness of the model itself.

Micali and Reyzin [14] have recently sought to address some of the deficiencies in current models of encryption by devising a corresponding model for *physically observable cryptography.* They proposed an extension to the complexity-theoretic model to embrace the notion that cryptographic algorithms are typically executed in a physical environment of a computer. In so doing, they sought to address the deficiencies that have arisen from the aforementioned side-channel attacks.

In this work I will address a different deficiency of current models, namely the failure to model conveyance of semantic meaning through the physical act of communication. Just as the physical act of computation has side effects that are usable to a cryptanalyst, so too does the physical act of communication produce features that can be used to the advantage of the cryptanalyst. I will develop several examples of the phenomenon to demonstrate my point, but one obvious example is the encoding of communication into packets for transmission on a packet-switched network. It has been observed by multiple others (see Section 6) that the packetization of communication often leaks information about the content. While it may not leak the exact contents of the packets themselves, it leaks knowledge about the communication, and provides a tempting target for cryptanalysis.

The goal of this work is perhaps more modest than that of Micali and Reyzin, since I do not put forward any reasonable model under which a secure cryptosystem could be constructed. Instead, I will advance the view that *the structure of the Internet as we know it may actually preclude the existence of any reasonable model for completely secure encryption.* Given the degree to which society has come to depend upon the

[*] Updates to this paper may appear at `http://mccurley.org/papers/traffic/`.

Internet, this is a startling possibility. Moreover, I will give examples to suggest that the phenomena is more general than just packet switched networks, and arises from many forms of communication via language.

While optimistic cryptologists should and will continue their quest for perfectly secure systems, there is no a priori reason why such a thing has to exist. Indeed, the entire framework of complexity-based security arguments would be radically changed if it turns out that P=NP, though it may also be argued that a polynomial separation between the capabilities of the legitimate user and the adversaries is sufficient for practical considerations. Moreover, the entire approach of complexity-theoretic security was an attempt to get around the limitation imposed by Shannon's result on perfect secrecy, and has proved to be remarkably effective in practice. It remains to be seen whether there is a similar approach that will mitigate the effects induced by the process of communication.

The point of view taken in this paper is partly historical and partly philosophical. An outline of the paper is as follows. In the next section we shall consider the definition of communication, after which we will present some examples of communication and how the process can leak knowledge. Following that we will propose a framework from which a partial security model can be constructed, without dwelling on the details. In fact, due to space and time constraints, I make no attempt describe a complete security model, but focus instead on the nature of the problem and why it may be impossible to construct such a model. I lay no claims on theorems regarding the possible existence or nonexistence of provably secure encryption. My hope is that this work will at least point the way toward better understanding of the underlying process of communication that we seek to model in the science of cryptology.

2 Mathematical Models of Encryption

During the last 60 years of mathematical research in cryptology, remarkable progress has been made in advancing cryptology to a science from what was once a black art. Most of the fundamental work has centered on the analysis of three models, namely information-theoretic security [19], complexity-based security [8], and quantum-theoretic security [2]. The goal of these is to construct a mathematical model of security, characterize the capabilities of an adversary, and (hopefully), provide a system that achieves some level of security under reasonable assumptions.

In both the information-theoretic model and the complexity-theoretic models of security, a secure cryptosystem is typically defined as a family of functions $E_k : M \to C$ that maps plaintexts $m \in M$ to ciphertexts $c \in C$. The family of encryption functions is indexed by the key $k \in K$ for some set of keys K. In Shannon's original formulation [19], an encryption system is said to have perfect secrecy if the adversary gains no more information about the plaintext from observing the ciphertext, i.e., $P_k(p|c) = P_k(p)$ for all keys $k \in K$. The major result that Shannon proved about this is that perfect secrecy requires that the key have as much entropy as the plaintext. This is often cited as a negative result, as it implies that substitution of secrecy of one piece of information (the plaintext) for another (the key) does not effectively result in any savings for the amount of secret information. This fact has motivated a lot of the research that has followed.

The example of the one-time pad is generally held up as the prototypical example of an encryption system that satisfies the perfect secrecy requirement, but in fact this holds only for messages that have constant length. The reason for this is obvious; the plaintext and the ciphertext are in fact the same length, so knowledge of the length of the ciphertext immediately reveals the length of the plaintext. While most theoreticians sweep this problem aside by simply assuming that all messages are the same size, I believe that this problem is in fact related to an important weakness in existing models.

In practice, the limitation of the one-time pad to message spaces in which all messages have the same length is at least as troublesome as the requirement for a large source of secure key bits. Moreover, the limitation is inherent to the definition of perfect secrecy, as is evidenced by the seminal observation of Chor and Kushilevitz [5] that it is impossible to construct an encryption system over a countably infinite message space that has information-theoretically perfect secrecy. Note that Shannon's original formulation incorporates an underlying probability distribution on plaintexts. This was reformulated in [5] by stating that for every pair of plaintexts p_1, p_2, $P(c|p_1) = P(c|p_2)$, or in other words, the probability of observing a given ciphertext is independent of the plaintext that generates it. Under this definition, they proved that there is no encryption system over a countably infinite message space that can achieve perfect secrecy. The implicit suggestion is that this is due to leakage of the length of the plaintext, and that this is unavoidable.

A primary driving force in the development of the complexity-theortic models of security was to address the fact that just because information about the plaintext would be present in the ciphertext need not compromise the plaintext, provided the adversary was constrained in their ability to compute the implicit information. A major breakthrough in this line of research was the construction by Goldwasser and Micali[10] of a *semantically secure* encryption system. A semantically secure encryption system is based on the notion of indistinguishability of ciphertexts; given two plaintexts it should be infeasible to distinguish which of them gave rise to a given ciphertext. A fundamental part of their construction was the realization that randomness is necessarily a part of any secure encryption system.

Unfortunately, it was proved by Oded Goldreich[9] that *a semantically secure encryption scheme must also leak information about the length of the plaintext.* A related problem lies at the heart of indistinguishability, namely that it does not address the issue of whether the eavesdropper can determine whether communication takes place, but only which message was sent. The mere fact that an eavesdropper observes the communication of bits from one party to another is in itself information, and knowledge about the number of bits is simply further leakage. The problem of leaking the size of the plaintext is often swept aside in mathematical treatments with the casual remark to simply pad or packetize all messages to be the same length. This approach was refuted in [9], but the problem has been largely ignored since then. It should be noted that Shannon [19] also chose not to address the problem of hiding the existence of communication, though he explicitly mentioned the distinction.

3 The Nature of Communication

In addressing the original problem of providing a reasonable definition for secure encryption of communication, it is prudent to consider what constitutes communication in the first place. In its purest form, communication is an amorphous concept, since the term is used to describe a variety of physical behaviors and other features in addition to the encoding of symbols. Moreover, it's not even clear what is being transferred in the act of communication. The problem of defining communication cuts very close to the often-cited DIKW hierarchy of data, information, knowledge, and wisdom. The definitions of such terms are hotly debated, lying on the boundary between philosophy, mathematics, and computer science. For a philosopher, knowledge is a topic in epistemology, and consists of thoughts that are true, believed, and justified. For a mathematician, knowledge is a concept in modal or temporal logic. For a computer scientist, knowledge represents the inference from and application of data and information, whereas information contains only answers to "who what where" questions. For followers of artificial intelligence, knowledge represents a degree of uncertainty. All of these points of view are probably relevant to the study of cryptology.

Consider the sentence "Why are you doing that?". At one level it can be thought of as a string of symbols (data). At a higher layer it consists of a sequence of words representing concepts (information). At an even higher layer, it has meaning as a question, though only within a context. The mere presence of the symbol '?' indicates that it is a question, but the mapping of interpretations from one layer to the next is seldom this transparent. The use of the term "that" indicates that the sentence only makes sense in a broader context, with reference either through physical proximity or through reference to an earlier information state.

Communication is often associated with action. If this is a sentence uttered from one person to another, then we probably should expect a response from the other party to shortly follow. If a response does follow, then we might expect it to be a response to the question. If the question is sent over a radio broadcast, then no such response is likely, since the channel does not support it. If the speaker of this sentence is waving their arms wildly then it probably has a different meaning than if the person is simply arching an eyebrow. All of these nuances can be considered elements of a model of communication, and all are potentially relevant to a cryptanalyst.

Unfortunately, models of secure encryption typically assume that the cryptanalyst is restricted to only the encoded symbols, or at best, to the concepts represented by the grouping of symbols. Cryptologic research has typically taken information theory as the the starting point for characterizing communication, starting from the seminal work of Claude Shannon. In this characterization, messages are emitted as blocks of symbols by the sender according to some known probability distribution, and that the problem is simply to conceal *which* of the possible messages was emitted. In practice, communication is much more complicated than this. Shannon's original model of communication was first published as a paper [17], in which he said:

> The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem.

Shannon's paper was republished the following year as part of a book, with introductory material by his coauthor, Warren Weaver [18]. The introductory material by Weaver alludes to the limitations of Shannon's definition for communication, and states that:

> In fact, two messages, one of which is heavily loaded with meaning and the other of which is pure nonsense, can be exactly equivalent, from the present viewpoint, as regards information. It is this, undoubtedly, that Shannon means when he says that "the semantic aspects of communication are irrelevant to the engineering aspects."

Weaver casts Shannon's theory as one layer of a more complex set of phenomena. Many of the advancements of the information age can be traced to the fundamental contributions of Shannon; this separation of meaning from encoding and transport is what allowed engineers to concentrate on a useful paradigm for technology, while ignoring the incredibly complex nuances of the underlying conveyance of concepts. Weaver identified three layers of problems in communication:

Level A. How accurately can the symbols of communication be transmitted? (the technical problem)

Level B. How precisely do the transmitted symbols convey the desired meaning? (the semantic problem)

Level C. How effectively does the received meaning affect conduct in the desired way? (the effectiveness problem)

Shannon and Weaver's separation of the communication problem into layers is analogous to the invention of written language, where complex human communication processes were reduced to a sequence of symbols on a page. The result was enormously powerful in influencing the nature of human communication because it eliminated the need for humans to be in physical proximity in order to communicate, but at the same time, something was lost in communication by the conversion to symbols. For example, a sentence that is spoken while waving arms wildly in the air has different semantic meaning than the same sentence that is uttered with arms crossed. Interpersonal communication often applies a secondary semantic interpretation or decoding of the communication that augments and corrects errors and omissions made from the spoken symbols. It should be noted that this nuance of definition for communication is also not limited to communication between humans. For example, it is easy to imagine how a computer will exhibit different characteristics of communication when it is in distress than when it is in a normal mode of operation.

4 The Nature of Cryptanalysis

The purpose of communication is to convey something, and in some cases that is merely to convey data. In this case, communication is thought of as stateless. In other cases, the goal is to convey something more, namely information that can be acted upon. In still other cases, the purpose is to create common knowledge out of knowledge. Our difficulty in defining encrypted communication is probably closely related to this confusion.

By breaking a communication system into layers, Shannon and Weaver were able to separate the problem of conveying ideas from that of conveying a symbolic representation of language. In Shannon's work, the semantic meaning of communication is separated from the problem of conveying it, since this was embodied in an encoding layer that takes place before and after the physical act of communication that was Shannon's focus. Unfortunately, from the point of view of a cryptanalyst, the semantic meaning of the underlying communication may be precisely what they are interested in, and the actual symbols used to convey the ideas may be of only peripheral interest. Cryptanalysis typically has a purpose, and the act of cryptanalysis is the gathering of actionable knowledge for this purpose. Thus while a cryptanalyst may be interested in recovering the credit card of a targeted person, they may also be interested in knowing what the person is buying, or of discovering their social preferences, or of simply knowing that a credit card was used. There is no direct way to quantify the range of semantic concepts that the cryptanalyst may be interested in, and in fact the information content of semantic information that may be derived from context of the communication can be arbitrarily large relative to the amount of information contained in the communication itself.

To illustrate the importance of semantic meaning in the process of cryptanalysis, consider the following questions that a cryptanalyst might ask about communication on the internet:

- What language is being spoken in a telephone call?
- Does Internet traffic contain VoIP or Skype traffic?
- Does Internet traffic use UDP or TCP?
- Is the same email being sent to multiple recipients?
- What is the nature of the relationships between the two parties in communication? Is one in command?
- What is the likelihood that a buy order will be issued in the next few seconds by a stock trader?

These are completely natural questions for a cryptanalyst to ask, and I claim that in each case there are plausible scenarios where the questions can be answered accurately with high probability, using observations of the physical act of communication.

It is tempting to define cryptanalysis as an attempt to create shared knowledge out of information. Unfortunately, it is completely unclear what falls within the domain of knowledge that is relevant to a given communication, since that requires us to characterize the goals of the eavesdropper relative to the two communicating parties who are his adversaries. It is almost certainly the case that any reasonable definition along these lines will need to take into account the state of knowledge of the eavesdropper before and after the communication, and the way in which it changed (either temporally, logically, or probabilistically). What is clear is that the current approach based only on information seems inadequate for accurately describing many situations.

5 The Use of Fragmentation in Communication

The original motivation for this work was to model the situation of two computers communicating privately over the Internet, and to understand the inherent limitations of using IPSEC to encrypt communication on the Internet. One of the fundamental features of Internet communication is that it is a packet switched network, in which the communication medium is shared between all parties connected to the network, and that communication is fragmented to enable congestion control and buffer management in intermediate routers. This feature of fragmentation of communication also arises in spoken and written language. Such "natural" language is typically composed of a sequence of distinct language elements (paragraphs, sentences, and words) that are themselves encoded into sequences of individual symbols or sounds.

To see why this the process of fragmentation is relevant to cryptanalysis, consider the following illustrative example. Suppose we are given the following fragment of encrypted text in which individual characters are encrypted but word breaks are exposed:

```
   # ####### ### ## ##### #### ###### ### ######## ######## #### #### ####
##### #### ## # #######
                                                                ###### ########
```

As a cryptanalyst we might begin by noticing that two of the words are only a single letter. If the original text is in English, we might expect these words to be either the letter “I” or the letter “A”. We might next notice that the text is arranged visually in a layout that is commonly used for quotations. The last line that is flush right might therefore be guessed to be a name, which greatly restricts the vocabulary. Knowing that this person is likely to be a famous person, we might be able to recover the most popular quotations of such people and apply a process of elimination. Even if the quotation was not in our list, we could apply basic knowledge of common sentence constructions to form a set of most likely candidates. If we hypothesis that the first letter is indeed ’A’, then we might further hypothesize that the next word is either an adjective or a noun. By knowing something about the context of the communication, we may form a hypothesis about the candidates for each word, and in the end arrive at a probability distribution on potential plaintexts that has a relatively low entropy from among all possible messages that fit the observed pattern.

There are several observations to be made from this simple example. First, our knowledge of word breaks provides a huge advantage for inferring the actual content of the message. Second, our knowledge of the underlying language and conventions for its usage assists us in identifying a few basic structures. Each of these factors interact with each other and increase our level of confidence in predicting the content of the message.

As another simple example, I took three versions of Tolstoy’s novel “Anna Karenina” written in English, French, and German. These should be more or less semantically the same message, with the only difference being that they are expressed in different languages. In order to test them to see if they could be distinguished from each other, I simply calculated the distribution of values of the lengths of the words. The result is shown in Figure 1. The data clearly shows a distinction between French and the other two languages.

Of course one should probably object to the relevance of this experiment since it’s hard to imagine a communication system that exposes word boundaries in language. On the other hand, nearly all existing text instant messaging protocols operate on the basis of buffering entire lines, which are often aligned with sentence boundaries. In this case the packets that would be sent would likely reveal the lengths of the sentences. The amount of information being leaked is less in this case, but for all we know it might still be possible to reliably distinguish between the topics of sports vs. travel, or whether the parties are male, or to make a good guess on the age of the sender. Moreover, there are numerous other examples where packetization of communication can reveal knowledge about the application.

6 Characteristics of Internet Communication

Two of the major characteristics that are present in Internet communications are *layering* and *packetization*. The principal of layering is ubiquitous in engineering of complex tasks such as networking, since it isolates the many different requirements from each other, and provides a layer of abstraction for one layer to address another. By separating the routing, transmission, ordering, buffering, physical device drivers, and error correction into different layers, it simplifies the maintenance of software systems, improves their reliability, and facilitates extensions to new technologies such as wireless.

The principle of packetization is probably the biggest single contributor to the success of the Internet, because the Internet is a shared network that provides transport for all parties who connect to it. By regulating and merging the flow of packets from different sources, the Internet provides congestion control and a degree of fairness in use of the shared network. It also facilitates buffering, error correction, and retransmission. Without packetization there would be no sharing, and by providing a shared network for multiple applications, the Internet greatly increased the efficiency of communication. Much of the value was realized due to the fact that different applications with different quality of service requirements can use the same underlying network infrastructure. Extreme examples of service requirements arise from voice over IP (VoIP) and HTTP. The primary quality of service requirement for VoIP is a high probability of delivery and low latency, since any interruption or delay of voice results in a poor user experience. By contrast, HTTP has a requirement for high throughput, since people would like to download more and more sophisticated pieces of content.

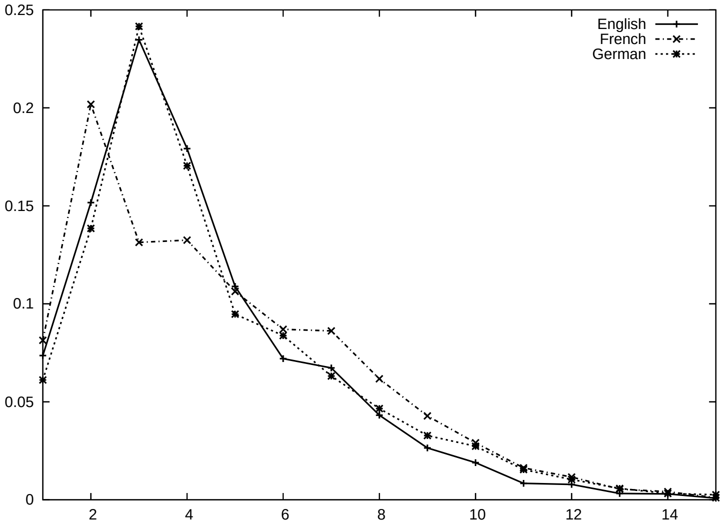

Fig. 1. Distribution of word lengths between three translations of the novel "Anna Karenina". Note that English and German have distinctly longer words on average than French, suggesting that it may be possible to distinguish French from the other two languages if word lengths are exposed through the communication process.

6.1 Cryptographic implications

The analogy of human natural language to Internet traffic is actually a very strong one, for the simple reason that word breaks are very much like packet boundaries, in that they reflect the semantics of the underlying communication. For example, in a persistent HTTP connection, the images embedded in a page are likely to be transmitted in packets that are separate from the packets containing the HTML page itself. In an interactive ssh session, a screen refresh event will often generate a packet containing as much of the refresh as will fit in a packet. This is due to the fact that within an application, individual `send()`s of information to the network are often broken into natural units of information that are defined by the application.

There are a number of complicated factors that determine whether an individual call by the program to `send()` generates a packet into the network, including the state of existing buffers, whether the application uses TCP or UDP, whether the operating system has properly implemented the TCP PUSH option, and whether the application chose to disable Nagle's algorithm on a TCP connection. In many cases it is still relatively easy to determine from the size and number of packets the number and size of `send()` calls in the application [11]. As a result, the "information breaks" of an application that correspond to word breaks or sentence breaks in natural language are often aligned to packet boundaries, and are therefore visible to an eavesdropper.

One feature of layering is that it provides the ability to address cryptographic requirements at the layer where it is convenient to do so. Examples include SSH at the application layer, SSL at the transport layer, IPSEC at the network layer, and WEP at the media layer. On the other hand, layering tends to introduce cryptographic weaknesses as well. For example, Bellovin [1] has observed that IP and TCP headers contain hints about the nature of the underlying traffic, and this largely results from layering, since quality of service is generally only implemented at the IP layer.

In addition to observations from headers, there are other signals present in packets from their timings, size, and patterns of traffic. For example, observations of SSH interactive login sessions were used to infer keystroke timings in [20], allowing them to mount an effective attack on passwords in SSH. A number of other examples were given by Bellovin [1].

The techniques of classifying traffic by characteristics that are not shielded by encryption have been developed by numerous authors, including Sun et. al. [21], Moore and Zuev [15], Zhang and Paxson [25], and Danezis [7, 6]. and Wright et. al. [24]. In spite of the increasing number of published attacks using characteristics of packet-switched networks, there has been very little discussion of this in the theoretical cryptography literature. Moreover, as was pointed out in section 2, current models of encryption simply avoid the problem of message size.

One application on the Internet that is rapidly gaining in popularity is voice over IP (VoIP). This protocol is extremely sensitive to latency and lost packets, so there are a number of optimizations and quality of service provisions for this service. Unfortunately some of these conflict with the security requirements of personal voice communication [23]. For example, VoIP voice packets are small (10-50 byte payload) and are therefore pretty easily recognized by their length. They are also likely to contain quality of service specifications in their headers. One interesting issue arises from a feature of VoIP that is designed to limit the bandwidth requirement for VoIP. In most phone conversations, only one end of the conversation will be talking at any given time. Hence it is only really important to send data in one direction most of the time, and in order to optimize bandwidth usage, VoIP supports something called silence suppression, where no packets are transmitted from the side that is silent. This feature has significant security implications, since this is precisely the kind of language break that was described in section 5!

7 Mathematial Models of Packetized Communication

Following up from the previous discussion, we can now derive some axioms that any model of packetized communication should follow in order to provide a meaningful model for cryptanalysis.

Axiom 1. A model of communication must include all sources and recipients of transmitted data. Consider for example a two-way communication between two people. A conversation may consist of questions, as well as responses to actions performed on the receipt of previous information. If we neglect to include these in our model, then we neglect a major source of information that is available to the eavesdropper.

Axiom 2. Communication is packetized. One way of looking at this is that communication has two states, namely when information is being transmitted and when it is not. Another way of saying this is that the sender is always sending; either real information or the null symbol, and the transmission of the null symbol is always detectable to the eavesdropper.

Axiom 3. Communication has state associated with it in both sender and receiver. This state changes as a result of receiving information.

Axiom 4. Communication has a temporal dimension, implying both an ordering and a distribution.

Axiom 5. Communication may be coupled to *observable* actions or states of the senders and recipients. In some cases traffic analysis may not be available to determine the source or destination of communication.

A natural model for a bidirectional channel is that of a pair of coupled Markov processes X_i, Y_i where X_i and Y_i are each dependent on $X_j, Y_j, j < i$. Here X_i, Y_i are ternary random variables taking on the values 0,1,null. The question of channel analysis is then to estimate the loss of information about X_i, Y_i when you are told when X_i, Y_i take on null values. More elaborate models would incorporate characteristics that may be observed about the aggregate of values, such as the notification tha a packet payload is beginning, or that a packet was fragmented, etc.

7.1 Keeping the channel full

The leakage of the length of the message may be regarded as a generalization of the fact that if an adversary observes communication taking place between two parties, then they gain some information that they were not previously in possession of. This fundamental problem lies within a class of attacks commonly referred to as “traffic analysis”. In practice this problem has been known for a very long time, and countermeasures are routinely used in modern link encryptors, by making sure that they always send information between sender and receiver, inserting dummy information if necessary [22]. By doing so, they seek to obscure the difference between actual communication and non-communication.

Unfortunately, the approach taken by link encryptors to "keep the channel full" is infeasible on the Internet, due to the requirement that the communication infrastructure serves the needs of multiple parties. In order for the Internet to operate efficiently and in an economically practical way, all parties must abstain from communicating except when they need to. One might ask how much additional bandwidth would be required in order for everyone to "keep the channel full". It has been observed empirically that the topology of the Internet connectivity graph has evolved as a sparse graph (e.g., see[16]), in which the degree distribution follows a power-law distribution. Thus in order to connect an Internet of n nodes, it appears that we require only $O(n)$ edges to provide a robust and scalable infrastructure for communication between potentially any pair of nodes. By contrast, if we adopt the link encryptor approach of masking the existence of communications by always communicating, we could potentially require $\binom{n}{2}$ edges in order for all n parties to be able to speak to each other. This is perhaps a pessimistic number, since the real number is the number of edges that a graph would require in order for there to be a collection of edge-disjoint paths between any bipartite matching of nodes in the graph. Of course even if the paths existed, we would still be left with the problem of finding them for routing purposes; this problem is unfortunately NP-complete. In other words, in order for the Internet to provide edge-disjoint paths that could be kept full between arbitrary matchings of nodes, a substantial increase in investment would be required.

8 Conclusions

The accumulated evidence of cryptanalysis through observation of communication points out that existing models of cryptographic security are lacking for at least two reasons. First, they fail to take into account the physical process of communication, in which the process of packetization is extremely important. It's almost certainly true that without packetization, the Internet could not have had the impact that it has. Yet at the same time, packetization has been seen to introduce numerous cryptographic weaknesses into communication, and there is currently no practical mathematical model to analyze the degree of weakness or within which we could prove anything about mitigating effects.

There is substantial doubt in the author's mind as to whether there is a reasonable balance that can be found between the quality of service demands of Internet applications and the goals of theoretical cryptography to provide an almost perfectly secure encryption methodology. It may turn out that it is inevitable that a cryptanalyst can attain some new knowledge from observing communication in some applications (notably those requiring low latency). If this is the case, then future research will be needed to define and quantify exactly how much knowledge will be leaked.

The second main point about theoretical models of cryptographic security is that they seem to overlook the distinction between knowledge and information. Shannon's achievement was to separate them so that communication engineering could proceed without the need to worry about conveyance of knowledge. Unfortunately, many cryptanalystic attacks take place at the knowledge layer of the DIKM hierarchy, and existing models fail to take this into account.

In many ways, the analysis of this paper may be regarded as being even more pessimistic than that of Shannon, since I have argued that the nature of Internet communication channels makes it inevitable that cryptanalysts will be able to gain knowledge from passive eavesdropping. I would be happy if I could be proved wrong.

References

[1] Steven M. Bellovin. Probable plaintext cryptanalysis of the IP security protocols. In *Proc. of the Symp. on Network and Distributed System Security*, pages 155–160, 1997.

[2] Charles Bennett, F. Bessette, Gilles Brassard, L. Salvail, and J. Smolin. Experimental quantum cryptography. *Journal of Cryptology*, 5:3–28, 1992.

[3] Eli Biham and Adi Shamir. Differential fault analysis of secret key cryptosystems. In *Advances in Cryptology, Proc. Crypto 1997*, Lecture Notes in Computer Science, pages 513–525. Springer-Verlag, 1997.

[4] Dan Boneh, Richard A. Demillo, and Richard J. Lipton. On the importance of checking cryptographic protocols for faults. In *Advances in Cryptology, Proc. Eurocrypt 1997*, volume 1233 of *Lecture Notes in Computer Science*, pages 37–51. Springer-Verlag, 1997.

[5] Benny Chor and Eyal Kushilevitz. Secret sharing over infinite domains. In *Proceedings of Crypto '89*, Lecture Notes in Computer Science, pages 299–306, Heidelberg, 1989. Springer-Verlag.

[6] George Danezis. Traffic analysis of the HTTP protocol over TLS. `http://homes.esat.kuleuven.be/~gdanezis/TLSanon.pdf`.

[7] George Danezis. Introducing traffic analysis: Attacks, defences and public policy issues, 2005. `http://homes.esat.kuleuven.be/~gdanezis/TAIntro.pdf`.

[8] Whitfield Diffie and Martin Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22:644–654, 1976.

[9] Oded Goldreich. A uniform-complexity treatment of encryption and zero-knowledge. *Journal of Cryptology*, 6:21–53, 1993.

[10] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28:270–299, 1984.

[11] Amit Klein. Detecting and preventing HTTP response splitting and HTTP request smuggling attacks at the TCP level. `http://www.securityfocus.com/archive/1/408135`.

[12] Paul Kocher. Cryptanalysis of Diffie-Hellman, RSA, DSS, and other cryptosystems using timing attacks. In *Advances in Cryptology, Proc. Crypto '95*, LNCS, pages 171–183. Springer-Verlag, 1995.

[13] Paul Kocher, Joshua Jaffe, and Benjamin Jun. Differential power analysis. In *Advances in Cryptology, Proc. Crypto '99*, LNCS, pages 388–397, Heidelberg, 1999. Springer-Verlag.

[14] Silvio Micali and Leonid Reyzin. Physically observable cryptography. In *Theory of Cryptography Conference*, volume 2951 of *LNCS*, pages 278–296. Springer, 2004.

[15] Andrew W. Moore and Denis Zuev. Internet traffic classification using Bayesian analysis techniques. In *SIGMETRICS '05*, pages 50–60, 2005.

[16] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.

[17] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423,623–656, 1948.

[18] C. E. Shannon and Warren Weaver. *The Mathematical Theory of Communication.* University of Illinois Press, 1949.

[19] Claude E. Shannon. Communication theory of secrecy systems. *Bell Systems Technical Journal*, pages 656–715, 1949.

[20] Dawn Xiaodong Song, David Wagner, and Xuqing Tian. Timing analysis of keystrokes and timing attacks on ssh. In *Proc. USENIX Security Symposium*, pages 337–352, Washington, D.C., 2001.

[21] Qixiang Sun, Daniel R. Simon, Yi-Min Wang, Wilf Russell, Venkata N. Padmanabhan, and Lili Qiu. Statistical identificatoin of encrypted web browsing traffic. In *Proc. IEEE Security and Privacy Symp.*, pages 19–30, 2002.

[22] V. L. Voydoc and Stephen Kent. Security mechanisms in high-level network pro-tocols. *ACM Computing Surveys*, pages 135–171, 1983.

[23] Thomas J. Walsh and Richard Kuhn. Challenges in security voice over IP. *IEEE Security and Privacy*, pages 44–49, May/June 2005.

[24] Charles Wright, Fabian Monrose, and Gerald M. Masson. HMM profiles for network traffic classification. In *ACM Conference on Computer and Communication Security*, pages 9–15, 2004.

[25] Y. Zhang and V. Paxson. Detecting stepping stones. In *Proc. 9th USENIX Security Symposium*, pages 171–184, 2000.