1 An information-theoretic approach to physical-layer security

A simple look at today's information and communication infrastructure is sufficient for one to appreciate the elegance of the layered networking architecture. As networks flourish worldwide, the fundamental problems of transmission, routing, resource allocation, end-to-end reliability, and congestion control are assigned to different layers of protocols, each with its own specific tools and network abstractions. However, the conceptual beauty of the layered protocol stack is not easily found when we turn our attention to the issue of network security. In the early days of the Internet, possibly because network access was very limited and tightly controlled, network security was not yet viewed as a primary concern for computer users and system administrators. This perception changed with the increase in network connections. Technical solutions, such as personnel access controls, password protection, and end-to-end encryption, were developed soon after. The steady growth in connectivity, fostered by the advent of electronic-commerce applications and the ubiquity of wireless communications, remains unhindered and has resulted in an unprecedented awareness of the importance of network security in all its guises.

The standard practice of adding authentication and encryption to the existing protocols at the various communication layers has led to what could be rightly classified as a patchwork of security mechanisms. Given that data security is so critically important, it is reasonable to argue that security measures should be implemented at *all* layers where this can be done in a cost-effective manner. Interestingly, one layer has remained almost ignored in this shift towards secure communication: the physical layer, which lies at the lowest end of the protocol stack and converts bits of information into modulated signals. The state of affairs described is all the more striking since randomness, generally perceived as a key element of secrecy systems, is abundantly available in the stochastic nature of the noise that is intrinsic to the physical communication channel. On account of this observation, this book is entirely devoted to an emerging paradigm: security technologies that are embedded at the *physical layer* of the protocol architecture, a segment of the system where little security exists today.

The absence of a comprehensive physical-layer security approach may be partly explained by invoking the way security issues are taught. A typical graduate course in cryptography and security often starts with a discussion of Shannon's information-theoretic notion of perfect secrecy, but information-theoretic security is quickly discarded and regarded as no more than a beautiful, yet unfeasible, theoretical construct. Such an exposition is designed to motivate the use of state-of-the-art encryption

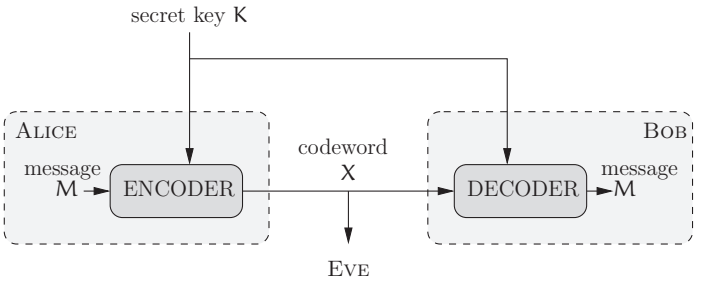

Figure 1.1 Shannon's model of a secrecy system.

algorithms, which are insensitive to the characteristics of the communication channel and rely on mathematical operations assumed to be hard to compute, such as prime factorization.

In this introductory chapter, we approach the subject in a different way. First, we give a bird's-eye view of the basic concepts of information-theoretic security and how they differ from classical cryptography. Then, we discuss in general terms some of the major achievements of information-theoretic security and give some examples of its potential to strengthen the security of the physical layer. The main idea is to exploit the randomness of noisy communication channels to guarantee that a malicious eavesdropper obtains no information about the sent messages: security is ensured not relative to a hard mathematical problem but by the physical uncertainty inherent to the noisy channel.

1.1 Shannon's perfect secrecy

Roughly speaking, the objective of secure communication is twofold; upon transmission of a message, the intended receivers should recover the message without errors while nobody else should acquire any information. This fundamental principle was formalized by Shannon in his 1949 paper [1], using the model of a secrecy system illustrated in Figure 1.1. A transmitter attempts to send a message M to a legitimate receiver by encoding it into a *codeword* X.[1] During transmission, the codeword is observed by an *eavesdropper* (called the enemy cryptanalyst in Shannon's original model) without any degradation, which corresponds to a worst-case scenario in which the communication channel is error-free. In real systems, where some form of noise is almost always present, this theoretical assumption corresponds to the existence of powerful error-correction mechanisms, which ensure that the message can be recovered with arbitrarily small probability of error. As is customary in cryptography, we often refer to the transmitter as "Alice," to the legitimate receiver as "Bob," and to the eavesdropper as "Eve."

In this worst-case scenario, the legitimate receiver must have some advantage over the eavesdropper, otherwise the latter would be able to recover the message M as well. The solution to this problem lies in the use of a secret key K, known only to the transmitter

[1] In cryptography, X is also called a cryptogram or ciphertext.

Table 1.1 Example of a one-time pad

Message	M	0	1	0	1	0	0	0	1	1	0	1
Key	K	1	0	0	1	1	0	0	0	1	0	1
Cryptogram	$\mathsf{X} = \mathsf{M} \oplus \mathsf{K}$	1	1	0	0	1	0	0	1	0	0	0

and the legitimate receiver. The codeword X is then obtained by computing a function of the message M and the secret key K.

Shannon formalized the notion of secrecy by quantifying the average uncertainty of the eavesdropper. In information-theoretic terms, messages and codewords are treated as random variables, and secrecy is measured in terms of the conditional entropy of the message given the codeword, denoted as $\mathbb{H}(\mathsf{M}|\mathsf{X})$. The quantity $\mathbb{H}(\mathsf{M}|\mathsf{X})$ is also called the eavesdropper's *equivocation*; *perfect secrecy* is achieved if the eavesdropper's equivocation equals the a-priori uncertainty one could have about the message, that is

$$\mathbb{H}(\mathsf{M}|\mathsf{X}) = \mathbb{H}(\mathsf{M})\,.$$

This equation implies that the codeword X is statistically independent of the message M. The absence of correlation ensures that there exists no algorithm that would allow the cryptanalyst to extract information about the message. We will see in Chapter 3 that perfect secrecy can be achieved only if $\mathbb{H}(\mathsf{K}) \geqslant \mathbb{H}(\mathsf{M})$; that is, the uncertainty about the key must be at least as large as the uncertainty about the message. In other words, we must have at least one secret bit for every bit of information contained in the message.

From an algorithmic perspective, perfect secrecy can be achieved by means of a simple procedure called a one-time pad (or Vernam's cipher), an example of which is shown in Table 1.1 for the case of a binary message and a binary key. The codeword is formed by computing the binary addition (XOR) of each message bit with a separate key bit. If the key bits are independent and uniformly distributed, it can be shown that the codeword is statistically independent of the message. To recover the message, the legitimate receiver need only add the codeword and the secret key. On the other hand, the eavesdropper does not have access to the key; therefore, from her perspective, every message is equally likely and she cannot do better than randomly guessing the message bits.

Although the one-time pad can achieve perfect secrecy with low complexity, its applicability is limited by the following requirements:

- the legitimate partners must generate and store long keys consisting of random bits;
- each key can be used only once (otherwise the cryptanalyst has a fair chance of discovering the key);
- the key must be shared over a secure channel.

To solve the problem of distributing long keys in a secure manner, we could be tempted to generate long pseudo-random sequences using a smaller seed. However, information theory shows that the uncertainty of the eavesdropper is upper bounded by the number of random key bits used. The smaller the key the greater the probability that the eavesdropper will succeed in extracting some information from the codeword. In this case,

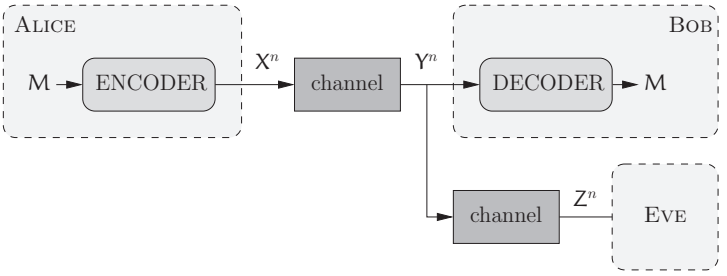

Figure 1.2 Wyner's wiretap channel model.

the only obstacle faced by the eavesdropper is computational complexity, which leads directly to the concept of computational security.

The aforementioned caveats regarding the one-time pad are arguably responsible for the skepticism with which security practitioners dismiss the usefulness of information-theoretic security. We shall now see that a closer look at the underlying communications model may actually yield the solution towards wider applicability.

1.2 Secure communication over noisy channels

As mentioned before, random noise is an intrinsic element of almost all physical communication channels. In an effort to understand the role of noise in the context of secure communications, Wyner introduced the *wiretap channel* model illustrated in Figure 1.2. The main differences between this approach and Shannon's original secrecy system are that

- the legitimate transmitter encodes a message M into a codeword X^n consisting of n symbols, which is sent over a noisy channel to the legitimate receiver;
- the eavesdropper observes a noisy version, denoted by Z^n, of the signal Y^n available at the receiver.

In addition, Wyner suggested a new definition for the secrecy condition. Instead of requiring the eavesdropper's equivocation to be exactly equal to the entropy of the message, we now ask for the *equivocation rate* $(1/n)\mathbb{H}(\mathsf{M}|\mathsf{Z}^n)$ to be arbitrarily close to the entropy rate of the message $(1/n)\mathbb{H}(\mathsf{M})$ for sufficiently large codeword length n. With this relaxed security constraint, it can be shown that there exist channel codes that *asymptotically* guarantee both an arbitrarily small probability of error at the intended receiver and secrecy. Such codes are colloquially known as *wiretap codes*. The maximum transmission rate that is achievable under these premises is called the *secrecy capacity*, and can be shown to be strictly positive whenever the eavesdropper's observation Z^n is "noisier" than Y^n.

In the seventies and eighties, the impact of Wyner's results was limited due to several important obstacles. First, practical code constructions for the wiretap channel were not available. Second, the wiretap channel model restricts the eavesdropper by assuming that

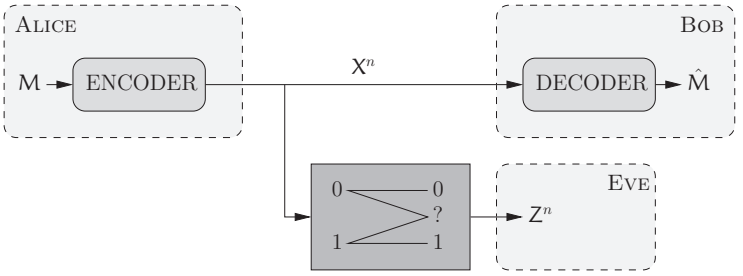

Figure 1.3 Communication over a binary erasure wiretap channel.

she suffers from more noise than is experienced by the legitimate receiver. In addition, soon after the notion of secrecy capacity appeared, information-theoretic security was overshadowed by Diffie and Hellman's seminal work on public-key cryptography, which relies on mathematical functions believed hard to compute and has dominated security research since then.

1.3 Channel coding for secrecy

Although the previous results on the secrecy capacity prove the existence of codes capable of guaranteeing reliable communication while satisfying a secrecy condition, it is not immediately clear how such codes can be constructed in practice. Consider the channel model illustrated in Figure 1.3, in which Alice wants to send one bit of information to Bob over an error-free channel while knowing that Eve's channel is a binary erasure channel, which erases an input symbol with probability ϵ. If Alice sends an uncoded bit, then Eve is able to obtain it correctly with probability $1-\epsilon$, leading to an equivocation equal to ϵ. It follows that, unless $\epsilon = 1$, Eve is able to obtain a non-trivial amount of information.

Alternatively, Alice could use an encoder that assigns one or more codewords to each of the two possible messages, 0 and 1. Suppose she takes all the binary sequences of length n and maps them in such a way that those with even parity correspond to $\mathrm{M} = 0$ and those with odd parity are assigned to $\mathrm{M} = 1$. If Bob receives one of these codewords over the error-free channel, he can obtain the correct message value by determining the parity of the received codeword. Eve, on the other hand, is left with an average of $n\epsilon$ erasures. As soon as one or more bits are erased, Eve loses her ability to estimate the parity of the binary sequence transmitted. This event happens with probability $1-(1-\epsilon)^n$ and it can be shown that

$$\mathbb{H}(\mathrm{M}|\mathrm{Z}^n) \geqslant 1-(1-\epsilon)^n,$$

which goes to unity as n tends to infinity. In others words, for sufficiently large codeword length, we get an equivocation that is arbitrarily close to the entropy of a message. The drawback of this coding scheme is that the transmission rate of $1/n$ goes to zero asymptotically with n as well. Alice and Bob can communicate securely by assigning

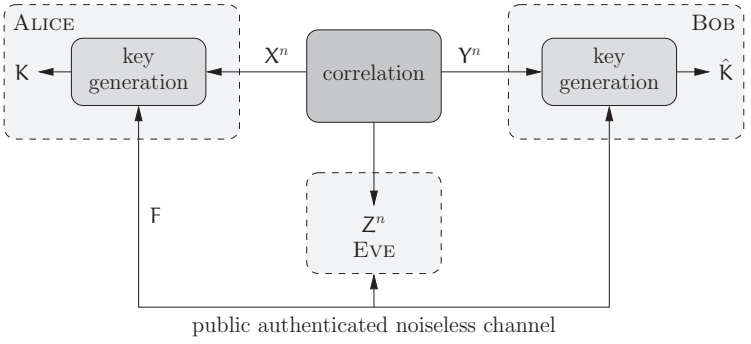

Figure 1.4 Secret-key agreement from correlated observations.

multiple codewords to the same message, but the secrecy achieved bears a price in terms of rate.

The intuition developed for the binary erasure wiretap channel should carry over to more general models. If Bob's channel induces fewer errors than Eve's, Bob should still be able to recover messages using a channel code; in contrast, Eve should be left with a list of possible codewords and messages. Asymptotic perfect secrecy can be achieved if this list covers the entire set of messages and their probability given that the received noisy codeword is roughly uniform. Unfortunately, to this day, practical wiretap code constructions are known for only a few specific channels.

1.4 Secret-key agreement from noisy observations

If Alice and Bob are willing to settle for generating a secret key instead of communicating a secret message straight away, then they can use the noisy channel to generate correlated random sequences and subsequently use an error-free communication channel to agree on a secret key. Such a situation is illustrated in Figure 1.4, in which Alice, Bob, and Eve obtain correlated observations X^n, Y^n, and Z^n, respectively; Alice and Bob then generate a key K on the basis of their respective observations and a set of messages F exchanged over the error-free channel. In the early nineties, Maurer and Ahlswede and Csiszár showed that, even if messages F are made available to the eavesdropper, Alice and Bob can generate a key oblivious to the eavesdropper such that $\mathbb{H}(K|Z^nF)$ is arbitrarily close to $\mathbb{H}(K)$. Provided that authentication is in place, granting Eve access rights to all feedback messages does not compromise security.

To gain some intuition for why public feedback is useful, consider an instance of Wyner's setup, in which the main channel and the eavesdropper's channel are binary symmetric channels. When Alice transmits a random symbol X over the main channel, Bob obtains $Y = X \oplus D$ and Eve observes $Z = X \oplus E$, where D and E are Bernoulli random variables that correspond to the noise added by the main channel and the eavesdropper's channel, respectively. Assume further that Bob's channel is noisier than Eve's, in the sense that $\mathbb{P}[D = 1] \geqslant \mathbb{P}[E = 1]$. Bob now uses the feedback channel in

the following manner. To send a symbol V, he adds the noisy observation received from the channel and sends $V \oplus Y = V \oplus X \oplus D$ over the public channel. Since Alice knows X, she can perform a simple binary addition in order to obtain $V \oplus D$. Eve, on the other hand, has only a noisy observation Z, and it can be shown that her optimal estimation of V is $V \oplus Y \oplus Z = V \oplus D \oplus E$. Thus, Alice and Bob effectively transform a wiretap scenario that is advantageous to Eve into a channel in which she suffers from more errors than do Alice and Bob.

From a practical perspective, the design of key-agreement schemes from correlated observations turns out to be a simpler problem than the construction of codes for the wiretap channel. In fact, a wiretap code needs to guarantee *simultaneously* reliable communication to the legitimate receiver and security against the eavesdropper. On the other hand, since a key does not carry any information in itself, the reliability and security constraints can be handled *separately*. For instance, Alice would first send error-correction data to Bob, in the form of parity bits, which would allow him to revert the bit flips caused by the noise in the channel. Even if the error-correcting bits are sent over the public channel, the fact that Eve's observation contains more errors than Bob's is sufficient to guarantee that she is unable to arrive at the same sequence as Alice and Bob. Alice and Bob would then use a well-chosen hash function to transform their common sequence of symbols into a much shorter key and, because of her errors, Eve is unable to predict the output of the hash. Finally, the key would be used as a one-time pad to ensure information-theoretic security.

1.5 Active attacks

Thus far, we have assumed that Eve is a passive eavesdropper, who wishes to extract as much information as possible from the signals traversing the channel. However, if she can afford the risk of being detected by the legitimate partners, she has a wide range of active attacks at her disposal. Eve could impersonate Alice or Bob to cause further confusion, intercept and forge messages that are sent over the noisy channels and the error-free public channel, or simply send jamming signals to perturb the communication.

Sender authentication is a tacit assumption in most contributions in the area of information-theoretic security. Except in special and rare instances of the wiretap scenario, a shared secret in the form of a small key is necessary to authenticate the first transmissions. Subsequent messages can be authenticated using new keys that can be generated at the physical layer using some of the methods in this book. Alternatively, if Alice and Bob are communicating over a wireless channel, then they can sometimes exploit the reciprocity of the channel to their advantage. The receiver can associate a certain channel impulse response with a certain transmitter and it is practically impossible for an adversary located at a different position to be able to generate a similar impulse response at the receiver.

With authentication in place, it is impossible for the attacker to impersonate the legitimate partners and to forge messages. However, the attacker may decide to obstruct the communication by means of jamming. This can be done in a blind manner by transmitting

noise, or in a more elaborate fashion exploiting all the available information on the codes and the signals used by the legitimate partners. It is worth pointing out that the use of jamming is not restricted to the active attackers. Cooperative jamming techniques, by which one or more legitimate transmitters send coded jamming signals to increase the confusion of the attacker, can be used effectively to increase the secrecy capacity in multi-user channels. Sophisticated signal processing, most notably through the use of multiple antennas, can also further enhance the aforementioned security benefits.

1.6 Physical-layer security and classical cryptography

There are many fundamental differences between the classical cryptographic primitives used at higher layers of the protocol stack and physical-layer security based on information-theoretic principles. It is therefore important to understand what these differences are and how they affect the choice of technology in a practical scenario.

Classical computational security uses public-key cryptography for authentication and secret-key distribution and symmetric encryption for the protection of transmitted data. The combination of state-of-the-art algorithms like RSA and the Advanced Encryption Standard (AES) is deemed secure for a large number of applications because so far no efficient attacks on public-key systems are publicly known. Many symmetric ciphers were broken in the past, but those that were compromised were consistently replaced by new algorithms, whose cryptanalysis is more difficult and requires more computational effort. Under the assumption that the attacker cannot break hard cryptographic primitives, it is possible to design systems that are secure with probability one. The technology is readily available and inexpensive.

However, there are also disadvantages to the computational model. The security of public-key cryptography is based on the conjecture that certain one-way functions are hard to invert, which remains unproven from a mathematical point of view. Computing power continues to increase at a very fast pace, such that brute-force attacks that were once deemed unfeasible are now within reach. Moreover, there are no precise metrics to compare the strengths of different ciphers in a rigorous way. In general, the security of a cryptographic protocol is measured by whether it survives a set of attacks or not. From the works of Shannon and Wyner, one concludes that the ruling cryptographic paradigm can never provide information-theoretic security, because the communication channel between the friendly parties and the eavesdropper is noiseless and the secrecy capacity is zero. Moreover, existing key-distribution schemes based on the computational model require a trusted third party as well as complex protocols and system architectures. If multiple keys are to be generated, it is usually possible to do so only from a single shared secret and at the price of reduced data protection.

The main advantages of physical-layer security under the information-theoretic security model come from the facts that no computational restrictions are placed on the eavesdropper and that very precise statements can be made about the information that is leaked to the eavesdropper as a function of the channel quality. Physical-layer security has already been realized in practice through quantum key distribution and, in theory,

suitably long codes can come exponentially close to perfect secrecy. The system architecture for security is basically the same as the one for communication. Instead of distributing keys, it is possible to generate on-the-fly as many secret keys as desired.

However, we must accept some disadvantages as well. First and foremost, information-theoretic security relies on average information measures. The system can be designed and tuned for a specific level of security, claiming for instance that with very high probability a block will be secure; however, it might not be possible to guarantee confidentiality with probability one. We are also forced to make assumptions about the communication channels that might not be accurate in practice. In most cases, one would make very conservative assumptions about the channels, which is likely to result in low secrecy capacities or low secret-key or message exchange rates. A few systems have been deployed, most notably for optical communication, but the technology is not very widely available and is still expensive.

In light of the brief comparisons above, it is likely that any deployment of a physical-layer security protocol in a classical system would be part of a layered security solution whereby confidentiality and authentication are provided at a number of different layers, each with a specific goal in mind. This modular approach is how virtually all systems are designed, so, in this context, physical-layer security provides an additional layer of security that does not yet exist in communication networks.

1.7 Outline of the rest of the book

The main objective of this book is to lay out the theoretical foundations of physical-layer security and to provide practical tools for implementing it in real systems. The different chapters cover essential theory and mathematical models for assessing physical-layer security and characterizing its fundamental limits, coding schemes for data security at the physical layer, and system aspects of physical-layer security.

Chapter 2 summarizes fundamental notions of information theory required in order to understand subsequent chapters. Our presentation emphasizes the mathematical tools and notions of particular relevance to physical-layer security.

Chapter 3 introduces the seminal results regarding secrecy capacity for communication channels, highlighting the mathematical techniques used in the derivations.

Chapter 4 focuses on the fundamental limits and methodologies of secret-key agreement, including the reconciliation of correlated sequences and how privacy amplification allows strong secrecy.

Chapter 5 discusses the fundamental limits of secure communication over Gaussian and wireless channels.

Chapter 6 covers some of the techniques used to achieve physical-layer security in practice, including the design of codes for wiretap channels as well as the construction of codes for secret-key agreement.

Chapter 7 addresses system issues related to the integration of physical-layer security in contemporary communications architectures and gives examples of practical applications.

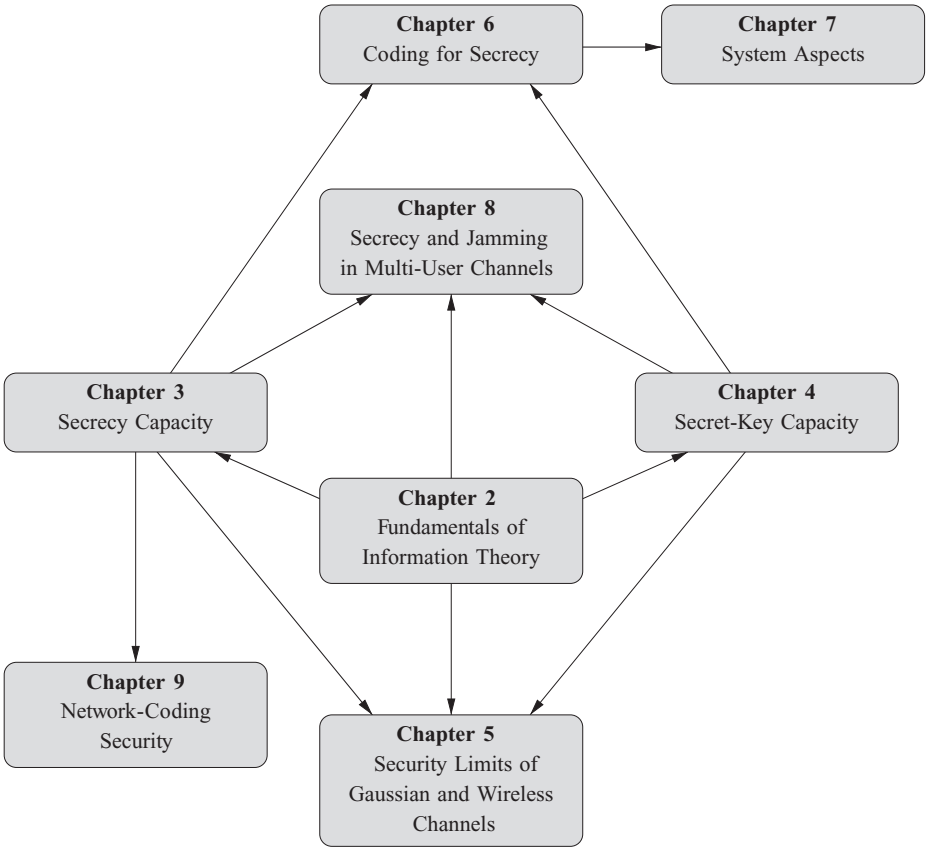

Figure 1.5 Dependences between chapters.

Chapter 8 discusses physical-layer security in multi-user systems and shows how the secrecy rates of multi-terminal networks can be increased through the appropriate use of feedback, cooperation, and jamming.

Chapter 9 deals with network-coding security. Although it is not necessarily implemented at the physical layer, network coding combines aspects of information and coding theory with close connections to information-theoretic security. By allowing intermediate nodes to mix different information flows through non-trivial operations, network coding offers a number of security challenges and opportunities.

The dependences between the chapters of the book are illustrated in Figure 1.5. The reader familiar with the tools and techniques of information theory can probably skip Chapter 2 and start at Chapter 3. The fundamental concepts and results of information-theoretic security are presented in Chapter 3 and Chapter 4 and are leveraged in Chapter 5 and Chapter 8 to study specific applications. Chapter 6, Chapter 7, and Chapter 9 rely on the notions introduced in earlier chapters but can be read independently.