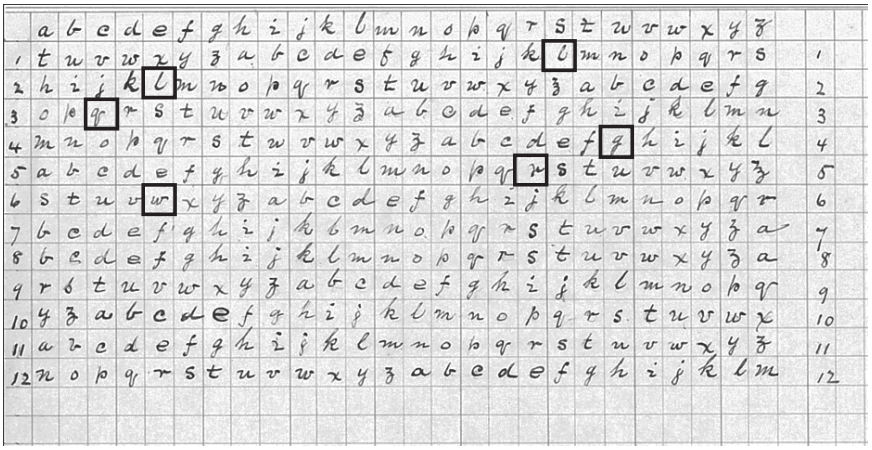
ciphertext letters, and the two occurrences of the ciphertext letter *l* represented different plaintext letters. This illustrates how the Vigenère cipher confounds simple frequency analysis, which was the main tool of cryptanalysts at the time. Although the idea may seem simple, the discovery of the Vigenère cipher is regarded as a fundamental advance in cryptography, and the method was considered to be unbreakable for hundreds of years.



Harvard University Archives.

FIGURE 5.5 A Vigenère cipher. The key, *thomasbbryan*, runs down the second column. Each row represents a Caesar cipher in which the shift amount is determined by a letter of the key. (Thomas B. Bryan was an attorney who used this code for communicating with a client, Gordon McKay, in 1894.)

CRYPTOGRAPHY AND HISTORY

Cryptography (code-making) and cryptanalysis (code-breaking) have been at the heart of many momentous events in human history. The intertwined stories of diplomacy, war, and coding technology are told beautifully in two books: *The Code-Breakers*, revised edition, by David Kahn (Scribner's, 1996) and *The Code Book* by Simon Singh (Anchor paperback, 2000).

Cryptographers use stock figures for describing encryption scenarios: Alice wants to send a message to Bob, and Eve is an adversary who may be eavesdropping.

Suppose Alice wants to send Bob a message (see Figure 5.6). The lock-and-key metaphor goes this way: Alice puts the message in a box and locks the box, using a key that only she and Bob possess. (Imagine that the lock on Alice's box is the kind that needs the key to lock it as well as to open it.) If Eve intercepts the

box in transit, she has no way to figure out what key to use to open it. When Bob receives the box, he uses his copy of the key to open it. As long as the key is kept secret, it doesn't matter that others can see that there is a box with something in it, and even what kind of lock is on the box. In the same way, even if an encrypted message comes with an announcement that it is encrypted using a Vigenère cipher, it will not be easy to decrypt, except by someone who has the key.

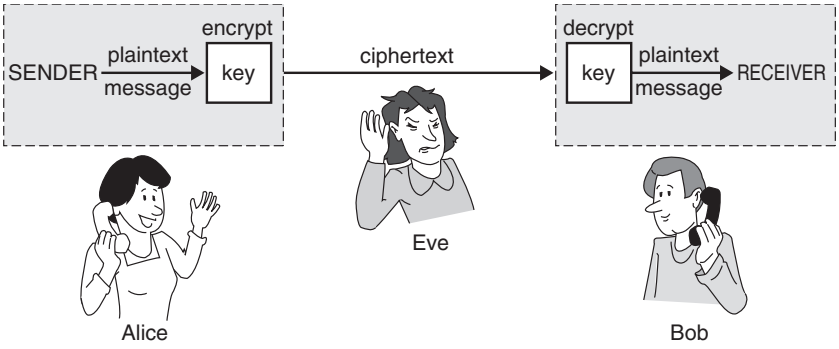


FIGURE 5.6 Standard cryptographic scenario. Alice wants to send a message to Bob. She encrypts it using a secret key. Bob decrypts it using his copy of the key. Eve is an eavesdropper. She intercepts the coded message in transit, and tries to decrypt it.

Or at least that's the idea. The Vigenère cipher was actually broken in the mid 1800s by the English mathematician Charles Babbage, who is now recognized as a founding figure in the field of computing. Babbage recognized that if someone could guess or otherwise deduce the length of the key, and hence the length of the cycle on which the Vigenère cipher was repeated, the problem was reduced to breaking several simple substitutions. He then used a brilliant extension of frequency analysis to discover the length of the key. Babbage never published his technique, perhaps at the request of British Intelligence. A Prussian Army officer, William Kasiski, independently figured out how to break the Vigenère code and published the method in 1863. The Vigenère cipher has been insecure ever since.

The sure way to beat this attack is to use a key that is as long as the plaintext, so that there are no repetitions. If we wanted to encrypt a message of length 100, we might use 100 Caesar ciphers in an arrangement like that of Figure 5.5, extended to 100 rows. Every table row would be used only once. A code like this is known as a *Vernam cipher*, after its World War I-era inventor, AT&T telegraph engineer Gilbert Vernam, and is more commonly referred to as a *one-time pad*.

The term "one-time pad" is based on a particular physical implementation of the cipher. Let's again imagine that Alice wants to get a message to Bob. Alice and Bob have identical pads of paper. Each page of the pad has a key written on it. Alice uses the top page to encrypt a message. When Bob receives it, he uses the top page of his pad to decrypt the message. Both Alice and Bob tear off and destroy the top page of the pad when they have used it. It is essential that the pages not be re-used, as doing so could create patterns like those exploited in cracking the Vigenère cipher.

One-time pads were used during the Second World War and the Cold War in the form of booklets filled with digits (see Figure 5.7). Governments still use one-time pads today for sensitive communications, with large amounts of keying material carefully generated and distributed on CDs or DVDs.

National Security Agency.

FIGURE 5.7 German one-time pad used for communication between Berlin and Saigon during the 1940s. Encrypted messages identified the page to be used in decryption. The cover warns, "Sheets of this encryption book that seem to be unused could contain codes for messages that are still on their way. They should be kept safe for the longest time a message might need for delivery."

A one-time pad, if used correctly, cannot be broken by cryptanalysis. There are simply no patterns to be found in the ciphertext. There is a deep relation between information theory and cryptography, which Shannon explored in 1949. (In fact, it was probably his wartime research on this sensitive subject that gave birth to his brilliant discoveries about communication in general.) Shannon proved mathematically what is obvious intuitively: The one-time

pad is, in principle, as good as it gets in cryptography. It is absolutely unbreakable—in theory.

But as Yogi Berra said, "In theory, there is no difference between theory and practice. In practice, there is." Good one-time pads are hard to produce. If the pad contains repetitions or other patterns, Shannon's proof that one-time pads are uncrackable no longer holds. More seriously, transmitting a pad between the parties without loss or interception is likely to be just as difficult as communicating the plaintext of the message itself without detection. Typically, the parties would share a pad ahead of time and hope to conceal it in their travels. Big pads are harder to conceal than small pads, however, so the temptation arises to re-use pages—the kiss of death for security.

The Soviet KGB fell victim to exactly this temptation, which led to the partial or complete decryption of over 3000 diplomatic and espionage messages by U.S. and British intelligence during the years 1942–1946. The National Security Agency's VENONA project, publicly revealed only in 1995, was responsible for exposing major KGB agents such as Klaus Fuchs and Kim Philby. The Soviet messages were doubly encrypted, using a one-time pad on top of other techniques; this made the code-breaking project enormously difficult. It was successful only because, as World War II wore on and material conditions deteriorated, the Soviets re-used the pads.

Because one-time pads are impractical, almost all encryption uses relatively short keys. Some methods are more secure than others, however. Computer programs that break Vigenère encryption are readily available on the Internet, and no professional would use a Vigenère cipher today. Today's sophisticated ciphers are the distant descendents of the old substitution methods. Rather than substituting message texts letter for letter, computers divide the ASCII-encoded plaintext message into blocks. They then transform the bits in the block according to some method that depends on a key. The key itself is a sequence of bits on which Alice and Bob must agree and keep secret from Eve. Unlike the Vigenère cipher, there are no known shortcuts for breaking these ciphers (or at least none known publicly). The best method to decrypt a ciphertext without knowing the secret key seems to be brute-force exhaustive search, trying all possible keys.

The amount of computation required to break a cipher by exhaustive search grows exponentially in the size of the key. Increasing the key length by one bit doubles the amount of work required to break the cipher, but only slightly increases the work required to encrypt and decrypt. This is what makes these ciphers so useful: Computers may keep getting faster—even at an exponential rate—but the work required to break the cipher can also be made to grow exponentially by picking longer and longer keys.