Although this is the most obvious way to identify spam, it's also the least reliable. Spammers learned long ago to leave common words out of their spams to avoid these types of filters. Often they intentionally misspell words or substitute numbers or symbols for letters, such as the numeral 0 for the letter *o,* or the symbol ! for the letter *l.*

The biggest problem with keyword checking is that it often leads to false positives. Friends and relatives might intentionally or inadvertently use any of the banned words in their emails. Sometimes, the banned words appear in the middle of otherwise completely innocent words. For example, if you list *Cialis* as a keyword that you want blocked, you'll also block the words *specialist* or *socialist.*

For these reasons, keyword filters are typically used only for the most obvious and offensive words and phrases, if they're used at all.

- **Bayesian analysis:** One of the most trusted forms of spam filtering is *Bayesian analysis,* which works by assuming that certain words occur more often in spam email than in other email. This sounds a lot like keyword checking, but Bayesian analysis is much more sophisticated than simple keyword checking. The Bayesian filter maintains an index of words that are likely to be encountered in spam emails. Each word in this index has a probability associated with it, and each word in the email being analyzed is looked up in this index to determine the overall probability of the email being spam. If the probability calculated from this index exceeds a certain threshold, the email is marked as spam.

  Here's where the magic of Bayesian analysis comes in: The index is self-learning, based on the user's actual email. Whenever the filter misidentifies an email, the user trains the filter by telling the filter that it was incorrect. The user typically does this by clicking a button labeled "This is spam" or "This is not spam." When the user clicks either of these buttons, the filter adjusts the probability associated with the words that led it to make the wrong conclusion. So, when the filter encounters a similar email in the future, it's more likely to make the correct determination.

- **Sender Policy Framework (SPF):** Surprisingly, SMTP (the Internet email protocol) has very poor built-in security. In particular, any email server can easily send email that claims to be from any domain. This makes it easy to forge the From address in an email. SPF lets you designate via DNS which specific email servers are allowed to send email from your domain. An antispam SPF filter works by looking up the sending email server against the SPF records in the DNS of the domain specified by the email's From address.

- **Blacklisting:** Another trusted form of spam filtering is a *blacklist* (also known as *blocklist*), which uses a list of known spammers to block email from sources that aren't trustworthy. There are two types of blacklists: private and public. A private blacklist is a list that you set up yourself to designate sources you don't want to accept email from. A public blacklist is a list that is maintained by a company or organization and is available for others to use.