Barracuda Bayesian
Barracuda Spam Firewall

The following is a set of instructions to help understand, and utilize the new Bayesian learning system in place on the Barracuda Spam Firewall in version 3.3 and higher of the Firmware. If you have any questions after reading this document, please call us at 408-342-5400 or email us at support@barracudanetworks.com.

What is Barracuda Bayesian and how do I get it?

The Barracuda Bayesian system is a replacement for the Bayesian system in use in the earlier versions of the Firmware. It is present in the 3.3 Firmware series and can be obtained by upgrading to version 3.3.01 or higher under **`Advanced->Firmware Update`** in the web interface.

The new Bayesian system provides many enhancements to the Bayesian system. These enhancements include the following:

- Ability to view user Bayesian statistics for a user account under **`Preferences->Spam Settings`** in the user's account of the web interface.
- Advanced Token Chaining (ATC) algorithm for higher context sensitivity in the Bayesian scoring process.
- Barracuda journaling process for support initiated replay of classified messages.
- Synchronization of classified messages (available in 3.3.01.008 and higher) within a cluster for learning on all machines.
- Multiple algorithms available for use in message scoring.
- Increased message learning, and message scoring, throughput.
- Header insertion containing Bayesian score for each recipient.
- Fallback to Global Bayesian database if the User Bayesian database is not fully trained.
- Less affected by various Bayesian poisoning tactics used in Spam messages.
- Only Internationalized Bayesian system available today

Will I lose my old Bayesian Database?

Yes, as part of the upgrade process to the 3.3 Firmware the old Bayesian Database will be reset and the new system put into its place. This is a one-time reset, and is a blessing in disguise for many customers.

All Bayesian systems rely on the fact that messages classified are not much different than new messages arriving. Over time however, the Spam messages change drastically and the Bayesian system – while initially able to compensate for the new format – gradually declines in its effectiveness. When this happens new classifications are needed to update the Bayesian database.

Another common occurrence with Bayesian databases is that, over time, messages containing various Bayesian Poisoning tactics are classified resulting in a database that contains conflicting information about what is Spam and what is Not-Spam. In this situation a reset is usually the best way to address this problem, as further classifications may not be effective depending on how many conflicted messages had been classified. Note: *The new Barracuda Bayesian system is less affected by Bayesian Poisoning tactics and therefore will not need to be reset if this situation were to occur*.

Although the old database is reset, the effort needed to start obtaining effective scoring with it is much lower than the version provided with the prior Firmware. The rest of this document will address the "Best Practice" training techniques for getting up and running with a quality Bayesian database.

Barracuda Bayesian Training - "Best Practices"

When training the Barracuda Bayesian system it is best to abide by the following principles for message classification:

1. Only classify the bare minimum needed to get the system scoring messages. This value is currently set to 200 Spam messages and 200 Not-Spam messages and is the required for both User Bayesian and Global databases. If 200 messages of each type are not classified for a User database it will fallback to the Global database and if the required count has not been reached for the Global database the Bayesian scoring will be bypassed.

2. After the Bayesian scoring has been activated begin reviewing messages in the message log area. This process involves reading a message and viewing the X-Barracuda-Bayes header that was applied. This header will show the scoring breakdown for each recipient in the message in the following format.
   - X-Barracuda-Bayes: `FLAG DATABASE PROBABILITY CONFIDENCE SCORE`
     - FLAG: Indicates whether a message was Spam or Not-Spam
     - DATABASE: Indicates whether a user or global database was used.
     - PROBABILITY: Chance the message is Spam.
     - CONFIDENCE: Chance the scoring was accurate based on tokens in use.

   The following are samples headers for a Spam message and for a Not-Spam Message:

   X-Barracuda-Bayes: SPAM GLOBAL 0.9997 1.0000 4.4864
   X-Barracuda-Bayes: INNOCENT GLOBAL 0.6219 1.0000 0.1348

3. If any messages are scored as Spam but should not have been, then classify those messages as Not-Spam. If any messages are incorrectly scored as Not-Spam, then classify those appropriately as well. Note: *The Bayesian score is added to the score from Spam Assassin and therefore the header should be viewed instead of relying on the final score for the message.*

4. Do **not** over-classify messages. A minimal amount of messages should be classified once scoring is being applied, and during each classify session some of each type should be classified to keep things balanced. Repeat steps 2 and 3 until the system is accurately scoring messages with your Bayesian database. **If scores are accurate, then do not classify the message.** Re-enforcement occurs with the scoring process and over-classification will often be harmful instead of beneficial. Over time, as the scores start to drift, classification of a small amount of messages will be needed to update your database with the new trends.

5. If you are unable to get the scoring to close to your desired range within 800 messages of each type being classified then consider a manual reset of the Bayesian database and start fresh. Inaccurate scores with a high amount of messages classified is usually an indication that the messages being classified are not unique enough as Spam or Not-Spam. Try to utilize the following guidelines if you are having trouble determining whether a message should be classified:

   - Read the entire message. Often times a quick glance is not sufficient.
   - If the message is Spam, but there is content that might overlap with your Business or normal legitimate email then do not classify the message.
   - If the message is Not-Spam, but it appears to be a newsletter or some other opt-in content then look into a Whitelist for the sender (at either the global or user level) instead of Bayesian training.
   - Avoid classifying messages with random paragraphs (used as a Bayesian Poisoning tactic). Although the new Barracuda Bayesian system offers some protection against these attacks, if you are having trouble obtaining consistent and accurate scoring it is usually best to defer classifying these messages.