SWAE v1.0
User Manual

(Server Program for Analysis and
Tagging of Electronic Mail)

San José, Costa Rica 2004

Contents

Introduction

This user manual is presented as a guide for those system administrators that wish to install and configure SWAE, the server program for analysis and tagging of electronic mail that is part of the system for detection and tagging of unsolicited or undesired e-mail or spam known as SpamWarn, which denotes the system or set of programs and not a particular program.

This manual is divided in three parts: Installation, Configuration and Use.

The first part explains topics such as the ways in which the user can get the SWAE program, and instructions about the installation of the software in the desired platform.

In the second part, we explain the configuration parameters of the server program and how they alter its behavior and the services each provides. The different values that these parameters are given, as well as the expected outcome of their use.

In the last section the SPAP protocol used in communicating with the server is described in detail. Within the Usage section, in the sub-section dedicated to Administration, we detail the steps required for daily administrative tasks such as user account creation and activation.

Typographic Conventions

The following typographic conventions are used in the scope of this document when denoting certain syntactic elements:

`Code` `Code`	This fixed-size font is used to denote file content samples, inputs or outputs of the system, commands, e-mail samples, etc. Some fragments or excerpts of dialogs with SWAE are highlighted with a yellow background color.
Italics	Italics are used to highlight some contents inside the text.
Bold	And bold to denote reserved words used within the text itself.

```
CLI> COM
SRV> ANSWER
CLI> COM PAR
SRV> Line 1
SRV> Line 2
SRV> .
CLI>
```

Dialogs between the client and the server are shown in this way, in which the client commands are indicated with `CLI>` and the server responses are indicated with `SRV>`. Each command can have zero parameters (as in `CLI> COM`), or one parameter (as in `CLI> COM PAR`). Dialogs are alternating, that is, every client command will generate a response and the client must wait until a complete server response is received before sending a new command. Furthermore, the server is forced to generate a response that in turn allows the client to know the result of the last command, either successful or resulting in an error condition. Some client commands and server responses are made up of several lines and the end of the command or the response is given by a line with the hexadecimal ASCII code sequence `<2E><0D><0A>`, meaning a period (`.`), followed by a carriage return (`CR`) and a line feed (`LF`).

Installation

In this section you will find the instructions needed to install SWAE in some of the platforms for which the binary form is distributed. This manual is probably outdated (unless you are seeing it directly from our website) and it may be possible that the instructions for a particular version of an operating system are not shown. If this is your case, please go to our website and get the latest version of this manual.

How to get a copy of a binary distribution

In http://www.spamwarn.com, you will find a copy of the server program for your particular platform. You need to check the version of your operating system and the type of computer in which you are planning to install it because binaries are compiled individually for each operating system/equipment pair.

If your operating system version or equipment is not shown in the lists, probably the program is still not available for that platform. You can get in touch with us at the addresses on the website to coordinate the generation of a binary version for your configuration.

Steps for installation in FreeBSD 4.2

To install SWAE in FreeBSD, you will need administrative privileges in the system so you will need to be logged on as **root**.

Logged in as **root**, and having a compressed copy of the binary package, follow the next steps:

Uncompress the file with the following command, where ** is the version of SWAE that you got:

```
# gunzip swae-**_freebsd-4.2_i386.tar.gz
```

Now unpack the files contained in the `.tar` archive:

```
# tar -xvf swae-**_freebsd-4.2_i386.tar
```

Change the current directory to the directory where SWAE's files will be found:

```
# cd swae-**_freebsd-4.2_i386
```

Create the directory where SWAE will store its working files. A good example is `/usr/home/spamwarn`. Also, set the access rights.

```
# mkdir /usr/home/spamwarn
# chmod 755 /usr/home/spamwarn
```

Copy the `common` directory from the current directory to the directory created above:

```
# cp common /usr/home/spamwarn
```

Now copy the `swae` file to the `/usr/local/bin` directory

```
# cp swae /usr/local/bin
```

and the `swae.conf` file to the `/usr/local/etc` directory

```
# cp swae.conf /usr/local/etc
```

Finally, in order to have SWAE started each time the operating system starts, add the following line to the local startup file, `/etc/rc.local`. The same line is used from a command prompt to start the server.

```
/usr/local/bin/swae
```

This concludes the installation in FreeBSD. Before starting the server, it is recommended that you visit the Configuration section to learn which are the default options and what is their meaning.

Steps for installation in cygwin 1.14

The installation of SWAE in cygwin 1.14 on Windows requires that you are logged on as the system Administrator because it is mandatory that you first install cygwin 1.14 on the system.

As a first step, you must obtain and install cygwin 1.14. The cygwin installers, as well as the steps to do this can be found at the website http://www.cygwin.com. The rest of the steps for the installation of SWAE assume that cygwin 1.14 is installed in the default directory, that is, `C:\cygwin`. Also, make sure that you have the following packages installed for cygwin: `gunzip` & `tar`.

Having cygwin installed, follow these steps.

Start cygwin. To do this, from Windows you can go to Start->Run and type the following command line:

```
C:\cygwin\cygwin.bat
```

Once cygwin starts and the console appears, copy SWAE's installation file from the location where it was downloaded to, for example, C:\Downloads. ** is the version of SWAE that you downloaded.

```
$ cp /cygdrive/c/downloads/swae-**_cygwin-1.14_i386.tar.gz .
```

uncompress the file with the following command:

```
$ gunzip swae-**_cygwin-1.14_i386.tar.gz
```

Now unpack the files within the `.tar` archive:

```
$ tar -xvf swae-**_cygwin-1.14_i386.tar
```

Change the current directory to the directory where SWAE's files are found.

```
$ cd swae-**_cygwin-1.14_i386
```

Create the directory where SWAE will store its working files. A good example is `/home/spamwarn`. Also, set the access rights.

```
$ mkdir /home/spamwarn
$ chmod 755 /home/spamwarn
```

Copy the `common` directory from the current directory to the directory created above:

```
$ cp common /usr/home/spamwarn
```

Now copy the `swae.exe` file to the `/usr/local/bin` directory

```
$ cp swae.exe /usr/local/bin
```

and the `swae.conf` file to the `/usr/local/etc` directory

```
$ cp swae.conf /usr/local/etc
```

finally, to have SWAE started each time the operating system starts, enter the `start.reg` file into the Windows registry by typing the following command:

```
$ regedit start.reg
```

and confirm the messages that appear.

This concludes the installation in cygwin over Windows. Before starting the server, it is recommended that you visit the Configuration section to learn which are the default options and what is their meaning.

Configuration

In this section the configurable options of SWAE found in the `/usr/local/etc/swae.conf` file are presented. They are shown with the accompanying default values that the options take if not specified. To change these values just edit the file and restart the server.

Parameter	Default	Meaning
`COLLMSGQUOTA`	`0` (zero)	This is the maximum amount of messages that will be stored in each user's collections. The first message over that amount is moved to the common collection that is used in calculating extra probabilities for all users. If there is more than one message over that amount, the first is kept and moved, and the rest discarded. The default value means that there is no limit in the amount of messages. This is not a good value to use because requires that the administrator manually checks the collections regularly so that disk space is not used completely. It also avoids the common collection from populating. A good value for this parameter may be between `300` and `500`.
`COLLSIZQUOTA` (`NOT IMPLEMENTED`)	`0` (zero)	This value is very similar to the previous but with the maximum amount of space that the collections will use. In practice, though, as messages tend to have an average size, both parameters will correlate and `COLLMSGQUOTA` will suffice.
`APPROACH`	`GARY`	This parameter may eventually be deprecated but in the meantime it means, to use Paul Graham's approach (`APPROACH=PAUL`) to combine probabilities using the Naive Bayes theorem, or the more advanced, Gary Robinson's approach (`APPROACH=GARY`), to combine probabilities using the Chi-Squared Probability Distribution. The default value means that the Robinson approach will be used.

Parameter	Default	Meaning
THRESHOLD	0.48	This is the probability value over which e-mail is considered spam and legitimate otherwise. When Paul's approach is used, a THRESHOLD near 0.9 should be used. When Gary's is used, a THRESHOLD near 0.48 works best.
MAXCONN (NOT IMPLEMENTED)	16	The maximum amount of connections that the server will simultaneously service. This is not implemented at the moment, which means that the server can service an unlimited amount of connections.
S_BGSTRENGTH & X_BGPROBABILITY	1 0.5	These parameters govern Gary Robinson's approach behavior and lack any meaning if APPROACH==PAUL. S_BGSTRENGTH can be described as the weight or strength that we want to give to our background information if nothing is known about a particular word found in an e-mail (i.e. if it has never been seen before). X_BGPROBABILITY is the probability that we want to give to the event that an e-mail having an unseen word is spam. Robinson points out that good initial values for these parameters are 1 and 0.5 respectively.
MAXMSGSIZE	65536	To avoid DoS or buffer overflow attacks, the maximum size that is allowed for messages. This is given in kilobytes. If no limit is desired, use a 0 (zero). The default value is 65536KB or 64MB. Although this may seem little to some, it is a pretty reasonable size, considering that almost no message will ever have that size. In any case, an attached file of 64MB or more should be sent by other means (e.g. FTP). NOTE: This value should be chosen carefully because the full message is stored in main memory while the program receives it.

Parameter	Default	Meaning
LISTENPORT	7923	As the name implies, the TCP port where the server program will listen for connections. The default value is 7923 which is an unused port in the range allowed for applications (i.e. 1024..65536). In the alphabetic phone code, 7923 corresponds to SWAE.
PATH	/usr/home/spamwarn	The location within the file system where SWAE will find the user collections. SWAE must have permission to write in this directory because it will frequently write in it. The default value is "/usr/home/spamwarn", which is in fact the path used as example in this manual for the installation in FreeBSD.
SYSLOGLEVEL	NOTICE	The cumulative level that the system will use to define which events generate an entry in the log and which don't. Selecting a level includes all the levels above it. For example, SYSLOGLEVEL=CRITICAL implies that messages in the CRITICAL, ALERT and EMERGENCY categories will all generate an entry in the log. Available levels are: -EMERGENCY : Messages meaning that the system is unusable -ALERT : Messages that require immediate action -CRITICAL : Messages that describe a critical condition -ERROR : Messages that describe an error -WARNING : Messages representing an emergency -NOTICE : Messages informing of normal but important events -INFO : Purely informational messages -DEBUG : Messages used in the debugging of the system

Use

In this section, the commands in the SPAP protocol are explained. This protocol is specifically designed for the SWAE server. SPAP means Spam Probability Analysis Protocol.

This protocol enables applications, such as the SWP-POP3 and SWP-SMTP proxies, to request services from SWAE for the analysis of e-mail received for users. With it, developers can create external programs that can benefit from SWAE's services without having the need to include them within the applications themselves.

The daily administrative tasks are also described here.

The SPAP Protocol

The protocol used by the clients connecting to the SWAE server will be called SPAP, or Spam Probability Analysis Protocol. It is a very simple protocol and only requires the client to select a mailbox from where the server will fetch the mail collections and then submit an email message for analysis against the mail collections selected by the mailbox. On return, the server will provide either a copy of the message with a tag on the Subject, if applicable, and three additional headers describing the message's ID and probability of being spam, or a summary line with the same information.

The protocol is then described as follows:

Connection

Upon connection, the user receives a greeting like this:

```
+ OK SpamWarn Analysis Engine (SWAE) v1.0 READY.
```

Mailbox or Account Selection

At this point, the server is ready to select the mailbox. The client has to do this before proceeding. It will do it by issuing the MBOX command. If the user's name is John Doe, the command will be something like:

```
MBOX jdoe
```

The server will check the requested mailbox and respond depending on the

following situations.

If the mailbox's collections are loaded and ready in memory:

```
+OK Mailbox ready.                                                    (*)
```

If the mailbox exists but its collections are not yet fully loaded:

```
+OK Mailbox loading...                                                (**)
```

If the mailbox is not active, meaning that it has not received an activation email from the service provider, it will indicate it like this:

```
+OK Mailbox ready (Not Active).                                       (*)
```

or

```
+OK Mailbox loading... (Not Active)                                   (**)
```

If the mailbox exists but its collections are not loaded and there is no available slot to load them:

```
-ERR No slot available, check license.
```

If the mailbox doesn't exist at all:

```
-ERR Mailbox not available, create with CRTE.
```

The CRTE command is explained later on.

When a mailbox's collections are being loaded, (i.e. `+OK Mailbox loading...`), the client has the option of waiting until the collections are completely loaded. It can find out whether they are, by using the MBOX command again:

```
MBOX jdoe
```

Where of course, “jdoe” would be the name of the previously requested mailbox, and the server will respond:

```
+OK Mailbox loading...     //if the mailbox's collections are still being loaded    (*)
+OK Mailbox ready.         //if the mailbox's collections are ready in memory      (**)
```

Again, the client decides to wait or not, but if the server responds with either * or ** above, either for the first or subsequent times, the client can already continue issuing the commands for the next steps in the protocol.

When the server is retrieving the email collections and loading their probabilities into the probabilities' tables, calculations are based on fewer information than is available when tables are fully loaded. If a request is

processed at that particular moment, attention to it may be given immediately at a penalty in the precision of the calculations, delayed to obtain more precise calculation of probabilities, or simply held until the tables are fully loaded with a (probably big) penalty in time. Depending on the needs of the client, either mode of operation will be available by issuing the `ATTN` command:

```
ATTN IMDT
```

Will give immediate attention to the request by using whatever contents are on the frequency tables at the moment of querying them. The server then replies:

```
+0K Attention to requests will be immediate resulting in less TIME but less PRECISI0N.
```

The command:

```
ATTN DLYD
```

(Not implemented) Will delay delay the attention of requests until at least half of the messages in the collections are loaded, and will invoke the response:

```
+0K Attention to requests will be delayed resulting in more TIME but more PRECISI0N.
```

And the command:

```
ATTN HELD
```

(Not implemented) Will delay any response to the request until the full email collections are fully loaded into the frequency tables. And the server replies:

```
+0K Attention to requests will be held resulting in most TIME but most PRECISI0N.
```

NOTE: The delayed mode is the default mode of operation. It is recommended though, that it is explicitly requested by the client.

Selection of Response Type

Next, the client selects what kind of response it wants from the server after the email is sent for analysis. It can command the server to reply with a revised copy of the original email with headers and tag added (if applicable), by sending:

```
RPLY RVSD
```

And the server will respond:

```
+OK Revised reply will be sent.
```

Or it can request that only the information about the analysis be returned, by issuing the command:

```
RPLY INFO
```

The server then responds:

```
+OK Information only will be sent.
```

NOTE: The default mode of operation, that is, the default reply of the server is by sending the information only. It is advisable to select either mode explicitly, to aid in clarity.

E-Mail Submission for Analysis

After selecting which form of reply is desired, the client is ready to request that an email be analyzed. It will do so by issuing the command:

```
DATA
```

To which the server will respond, much like in the SMTP protocol:

```
+OK Go ahead, end with <.>+<CR>+<LF> on a clean line.
```

But if the client has not yet selected a mailbox, the server has to respond:

```
-ERR No mailbox selected.
```

If everything is ok at this moment, the client will start sending the email to the server and will end it with a period on a line by itself, followed by the sequence <CR>+<LF> or (#0D #0A in ASCII). For example:

```
Return-Path: <artblxdonalddr@best-buyz.com>
Delivered-To: jdoe@spamwarn.com
Received: (qmail 15616 invoked by uid 508); 15 Jun 2003 20:49:43 -0000
Received: from unknown (HELO m101.best-buyz.com) (64.119.221.101)
  by 216.218.184.245 with SMTP; 15 Jun 2003 20:49:43 -0000
To: jdoe@spamwarn.com
Date: Sun, 15 Jun 2003 13:51:14 -0800
Message-ID: <1055710274.7693@m101.best-buyz.com>
X-Mailer: Mutt/1.3.17i
From: "perfectly legal" <cshwywDonalddr@best-buyz.com>
Reply-To: "perfectly legal" <dukmwyDonalddr@best-buyz.com>
Subject: Please keep this secret
Content-Type: text/html
MIME-Version: 1.0

If you have Digital Cable...

you gotta see this amazing little tool.

Follow this link to go to the web site
```

```
http://www.duanrui.com/digitool/index.php?RepID=FIVE

agtgpo. zfpuyntre^hygenargjbek(pbz .cimeqjdatatracking.
.
```

Server Response

Once the final period has been received, the server will analyze the email and report the findings to the client if the mailbox is active. If it is not, it will just return an exact copy of the original email with a legend at the bottom specifying that the mailbox is not active, like this

```
This mailbox is being analyzed by SpamWarn but results are not being shown because
it has not been activated.  Please activate this mailbox at http://www.spamwarn.com
```

or a single line if the client selected `RPLY INF0`:

```
+0K SWID: PLEASE-ACTIVATE-THIS-MAILB0X-AT-WWW.SPAMWARN.C0M SWP: ?.????????
```

But for active mailboxes, if the server was requested to send a revised (`RPLY RVSD`) version of the email, it will return an email like the one below which is essentially the same message but with “SPAM” added to the subject (if applicable), and the “Message-ID”, “X-SpamWarn” and “X-SpamWarnID” headers added to the headers the message already had.

```
Return-Path: <artblxdonalddr@best-buyz.com>
Delivered-To: jdoe@spamwarn.com
Received: (qmail 15616 invoked by uid 508); 15 Jun 2003 20:49:43 -0000
Received: from unknown (HEL0 m101.best-buyz.com) (64.119.221.101)
  by 216.218.184.245 with SMTP; 15 Jun 2003 20:49:43 -0000
To: jdoe@spamwarn.com
Date: Sun, 15 Jun 2003 13:51:14 -0800
Message-ID: <2004-03-25-11-40-38-241-A36E90C0-jdoe>
X-Mailer: Mutt/1.3.17i
From: "perfectly legal" <cshwywDonalddr@best-buyz.com>
Reply-To: "perfectly legal" <dukmwyDonalddr@best-buyz.com>
Subject: SPAM: Please keep this secret
Content-Type: text/html
MIME-Version: 1.0
X-SpamWarn: SPAM 0.99994121
X-SpamWarnID: 2004-03-25-11-40-38-241-A36E90C0-jdoe

If you have Digital Cable...

you gotta see this amazing little tool.

Follow this link to go to the web site
http://www.duanrui.com/digitool/index.php?RepID=FIVE

agtgpo. zfpuyntre^hygenargjbek(pbz .cimeqjdatatracking.
.
```

Or, if the server was asked to send only the information on what the server found about the message (`RPLY INF0`), it will reply like this:

```
+0K SWID: 2004-03-25-11-40-38-241-A36E90C0-jdoe SWP: SPAM 0.99994121
```

Which means that the message has been assigned SpamWarn ID `2004-03-25-11-40-38-241-A36E90C0-jdoe` and a SpamWarn probability of `0.99994121` which means that it is most probably a SPAM message.

Previously Seen Messages

If the message submitted already has a X-SpamWarnID header (an this is the first thing SWAE must check), it means that most probably it has already been analyzed by SWAE and sent to the user. The user has decided that the message is not correctly tagged (either tagged as spam when it is not, or not tagged when it should), and has sent the message back to himself so that SWAE changes the previous tag.

When a message with a X-SpamWarnID header is received, SWAE will extract the value of the header, try to find the message within the collections, move it to the other collection and update the frequency tables. If the client has requested a revised reply, it will receive a confirmation message like this:

```
Return-Path: <"SWAE">
Delivered-To: jdoe@spamwarn.com
Received: (qmail 7005 invoked from network); 2 Apr 2004 19:08:50 -0000
Received: from spamwarn.ultranetworx.com (HELO spamwarn) (192.168.0.22)
  by walnut.phpwebhosting.com with SMTP; 2 Apr 2004 19:08:50 -0000
From: SpamWarn Server
To: <jdoe@spamwarn.com >
Subject: Message Tag Correction: GOOD: Antonio Lam, Legislador para el 8-8
Sender: SpamWarn Server
Mime-Version: 1.0
Content-Type: text/plain; charset="iso-8859-1"
Date: Fri, 2 Apr 2004 14:08:53 -0500
X-Mailer: SpamWarn Analysis Engine v1.0

Message from SpamWarn Server

Your message with SWID: 2004-03-25-11-40-38-241-A36E90C0-jdoe originally marked as
SPAM: 0.98235688, has been correctly recognized as GOOD: 0.50000000.
.
```

If the client has requested an information only reply, the server will reply with:

```
+OK SWID: 2004-03-25-11-40-38-241-A36E90C0-jdoe SWP-CORRECTED: GOOD: 0.00000000.
```

or, when re tagging as spam:

```
+OK SWID: 2004-03-25-11-40-38-241-A36E90C0-jdoe SWP-CORRECTED: SPAM: 1.00000000.
```

The CRTE Command

Another command to explain in this protocol is the CRTE command which can be used to create a mailbox if it does not exist. It is invoked with the name of the mailbox to be created, for example the user mailbox for Groucho Marx:

```
CRTE gmarx
```

To which the server responds:

```
+OK Mailbox ready (Not Active).
```

Which means that the mailbox was successfully created and it is ready to receive requests. Note also that the message also indicates that it is not active. To activate it, the steps described in the section about **Activation** on this chapter must be performed.

If for any reason, including that the mailbox already exists, that there would be no slot left to load it or access was denied when trying to create the mailbox on disk:

```
-ERR Mailbox not created.
```

Session End

There is one last command, QUIT:

```
QUIT
```

To which, as expected, the server responds with:

```
+OK Goodbye.
```

and then closes the connection to the client. It leaves any mailbox's collection's frequency tables loaded in a slot in the server's memory to speed subsequent queries.

Miscellaneous Notes

This protocol describes a service whose connections are fully reusable, meaning that once a mailbox is selected, it can receive requests, service them change parameters or another mailbox can be selected at any time (except when a DATA command has been sent because the server is not

servicing commands at that moment). Dialogs such as this are then possible:

```
MBOX jdoe
+OK Mailbox ready.
MBOX gmarx
+OK Mailbox ready.
```

Switching mailboxes is, as is selecting them for the first time, possible only if the mailbox exists and a slot for it has been allocated.

If a mailbox is ready for use, and another MBOX command is issued but the operation cannot be completed, SWAE defaults to the initial state, in which no mailbox is selected.

However, other commands, such as `ATTN` and `RPLY` are only valid once a mailbox has been selected. Otherwise, they will return an error such as:

```
-ERR No mailbox selected.
```

Summary

Summarizing, a usual SPAP dialog would have three basic stages:

1- Mailbox selection
2- Operation mode selection
3- Mail analysis request

For example a minimal dialog would be:

```
+OK SpamWarn Analysis Engine (SWAE) v1.0 READY.
MBOX jdoe
+OK Mailbox ready.
RPLY INFO
+OK Information only will be sent.
DATA
+OK Go ahead, end with <.>+<CR>+<LF> on a clean line.
Return-Path: <artblxdonalddr@best-buyz.com>
Delivered-To: jdoe@spamwarn.com
Received: (qmail 15616 invoked by uid 508); 15 Jun 2003 20:49:43 -0000
Received: from unknown (HELO m101.best-buyz.com) (64.119.221.101)
  by 216.218.184.245 with SMTP; 15 Jun 2003 20:49:43 -0000
To: jdoe@spamwarn.com
Date: Sun, 15 Jun 2003 13:51:14 -0800
Message-ID: <10557102274.7693@m101.best-buyz.com>
X-Mailer: Mutt/1.3.17i
From: "perfectly legal" <cshwywDonalddr@best-buyz.com>
Reply-To: "perfectly legal" <dukmwyDonalddr@best-buyz.com>
Subject: Please keep this secret
Content-Type: text/html
MIME-Version: 1.0

If you have Digital Cable...

you gotta see this amazing little tool.
```

```
Follow this link to go to the web site
http://www.duanrui.com/digitool/index.php?RepID=FIVE

agtgpo. zfpuyntre^hygenargjbek(pbz .cimeqjdatatracking.
.
+OK SWID: 2004-03-25-11-40-38-241-A36E90C0-jdoe SWP: SPAM 0.99994121
QUIT
+OK Goodbye.
```

Another example of a valid client/server dialog of a more human nature:

```
+OK SpamWarn Analysis Engine (SWAE) v1.0 READY.
DATA
-ERR No mailbox selected.
MBOX jdoe
+OK Mailbox loading...
MBOX jdoe
+OK Mailbox loading...
MBOX jdoe
+OK Mailbox ready.
RPLY RVSD
+OK Revised reply will be sent.
DATA
+OK Go ahead, end with <.>+<CR>+<LF> on a clean line.
Return-Path: <artblxdonalddr@best-buyz.com>
Delivered-To: jdoe@spamwarn.com
Received: (qmail 15616 invoked by uid 508); 15 Jun 2003 20:49:43 -0000
Received: from unknown (HELO m101.best-buyz.com) (64.119.221.101)
  by 216.218.184.245 with SMTP; 15 Jun 2003 20:49:43 -0000
To: jdoe@spamwarn.com
Date: Sun, 15 Jun 2003 13:51:14 -0800
Message-ID: <10557102274.7693@m101.best-buyz.com>
X-Mailer: Mutt/1.3.17i
From: "perfectly legal" <cshwywDonalddr@best-buyz.com>
Reply-To: "perfectly legal" <dukmwyDonalddr@best-buyz.com>
Subject: Please keep this secret
Content-Type: text/html
MIME-Version: 1.0

If you have Digital Cable...

you gotta see this amazing little tool.

Follow this link to go to the web site
http://www.duanrui.com/digitool/index.php?RepID=FIVE

agtgpo. zfpuyntre^hygenargjbek(pbz .cimeqjdatatracking.
.
Return-Path: <artblxdonalddr@best-buyz.com>
Delivered-To: jdoe@spamwarn.com
Received: (qmail 15616 invoked by uid 508); 15 Jun 2003 20:49:43 -0000
Received: from unknown (HELO m101.best-buyz.com) (64.119.221.101)
  by 216.218.184.245 with SMTP; 15 Jun 2003 20:49:43 -0000
To: jdoe@spamwarn.com
Date: Sun, 15 Jun 2003 13:51:14 -0800
Message-ID: <10557102274.7693@m101.best-buyz.com>
X-Mailer: Mutt/1.3.17i
From: "perfectly legal" <cshwywDonalddr@best-buyz.com>
Reply-To: "perfectly legal" <dukmwyDonalddr@best-buyz.com>
Subject: SPAM: Please keep this secret
Content-Type: text/html
MIME-Version: 1.0
X-SpamWarn: SPAM 0.99994121
X-SpamWarnID: 2004-03-25-11-40-38-241-A36E90C0-jdoe
```

```
If you have Digital Cable...
you gotta see this amazing little tool.

Follow this link to go to the web site
http://www.duanrui.com/digitool/index.php?RepID=FIVE

agtgpo. zfpuyntre^hygenargjbek(pbz .cimeqjdatatracking.
.
RPLY INF0
+0K Information only will be sent
DATA
+0K Go ahead, end with <.>+<CR>+<LF> on a clean line.
Return-Path: <artblxdonalddr@best-buyz.com>
Delivered-To: jdoe@spamwarn.com
Received: (qmail 15616 invoked by uid 508); 15 Jun 2003 20:49:43 -0000
Received: from unknown (HEL0 m101.best-buyz.com) (64.119.221.101)
  by 216.218.184.245 with SMTP; 15 Jun 2003 20:49:43 -0000
To: jdoe@spamwarn.com
Date: Sun, 15 Jun 2003 13:51:14 -0800
Message-ID: <10557l0274.7693@m101.best-buyz.com>
X-Mailer: Mutt/1.3.17i
From: "perfectly legal" <cshwywDonalddr@best-buyz.com>
Reply-To: "perfectly legal" <dukmwyDonalddr@best-buyz.com>
Subject: Please keep this secret
Content-Type: text/html
MIME-Version: 1.0

If you have Digital Cable...
you gotta see this amazing little tool.

Follow this link to go to the web site
http://www.duanrui.com/digitool/index.php?RepID=FIVE

agtgpo. zfpuyntre^hygenargjbek(pbz .cimeqjdatatracking.
.
+0K SWID: 2004-03-25-11-40-38-241-A36E90C0-jdoe SWP: SPAM 0.99994121
MB0X otheruser
-ERR Mailbox does not exist.
CRTE otheruser
+0K Mailbox created.
MB0X otheruser
+0K Mailbox ready.
MB0X yetanotheruser
+0K Mailbox ready.
RPLY RVSD
+0K Revised reply will be sent
DATA
+0K Go ahead, end with <.>+<CR>+<LF> on a clean line.
...
QUIT
+0K Goodbye.
```

Administration

SWAE does not require complex administrative routines. In fact, the server is to a high degree self-maintained. For example, the server is in charge of keeping a constant amount of messages in the collections, of refreshing them daily and of renovating them constantly with the new messages that arrive for each user. It also maintains a common collection that is added to each user's collections so that the amount of messages in them doubles and with that, the quality of the estimates, but more importantly, allows that a new user benefits from better quality of the estimates even when his account has been recently created and does not have enough messages available for training.

Nevertheless, there is a particular task which is responsibility of the administrator and that the server cannot perform on its own and that is the creation and activation of the mailboxes or user accounts.

It is required that the administrator manually connects to the server via a TELNET session and requests the creation of the account, for example, for the user Harold Lloyd:

```
CRTE hlloyd
+ OK Mailbox ready (Not Active).
```

Note again, that even though the mailbox has been created, it is not active yet.

Activation is requested by visiting the website http://www.spamwarn.com, where the administrator enters the e-mail address that he/she wants to activate and the website will automatically send a special activation e-mail to SWAE. This activation has an expiration period, usually of 45 days but is renewed every 30 days.

Another way in which an account is activated is the free trial period, in which the activation is good for 15 days and is granted only once when the account is created on the website.

Upon reception of the e-mail, SWAE gathers the required information to activate the account and notifies the user with an e-mail similar to this:

```
From: SpamWarn Server
To: jdoe@spamwarn.com
Subject: Activation Received
Sender: SpamWarn Server
Mime-Version: 1.0
Content-Type: text/plain; charset="iso-8859-1"
Date: 2004-11-10 14:54:22 +0600
X-Mailer: SpamWarn Analysis Engine v1.0

Message from SpamWarn Server
```

```
An activation message has been received for your account for 45 days.
```

Notes regarding account names

It is possible for the administrator to create accounts with special characters, for example, it is possible to create an account with the name “test 1, 2, 3”.

```
CRTE test 1, 2, 3.
```

Or even with the name ' “my account” ', including the quotes, like this:

```
CRTE "my account"
```

These names are allowed because internally, SWAE converts names such as “test 1, 2, 3.” to their URL-encoded representation like this “`test%201%2C%202%2C%203%2E`”. Besides convenient, this protects the server against the possibility of attack by directory traversal, for example if a command like this was executed:

```
CRTE ..\..\..\hackerdir
```

Which could probably cause the system to create the folder `hackerdir` at the root of the file system.

Although in theory, and for certain applications there is no restriction regarding the name given to an account, in practice there are some exceptions, particularly when using SWAE in conjunction with SWP-POP3 or SWP-SMTP.

When SWAE is used with the POP3 proxy, SWP-POP3, the administrator has to create accounts which are named identically to the account name in the POP3 server to which the proxy will connect.

In other words, if the user Albert Einstein has the e-mail account “aeinstein@geniuses.com”, the account in the POP3 server is probably “`aeinstein`”, and this will be the name of the user account in SWAE.

In some special cases, for example in environments with multiple virtual domains, it may be that Albert Einstein's account in the POP3 server is something like “`aeinstein%geniuses.com`”, which differentiates it from the account “`aeinstein%notsogeniuses.com`” in the same server.

When SWAE is used with the SMTP proxy, SWP-SMTP, a similar case arises. In SMTP users are identified by their full e-mail address, so in SWAE the account or mailbox name must be the full e-mail address.

In Albert Einstein's case, the SWAE account would no longer be “`aeinstein`” but “`aeinstein@geniuses.com`”

Conclusions

As it can be seen, the e-mail analysis server is very simple to use and allows its adaptation to different work environments. From proprietary applications where the client program is created by the development team of the company, to its use with the prepackaged proxies of the SpamWarn system, since it provides a well-defined protocol to that effect.

SWAE represents a step forward in spam avoidance technology because it offers a clear and concise interface, is highly configurable and requires little administration.