

Re: Non English Spam

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/questions/2006-10/msg01044.html>

- *From:* "Ted Mittelstaedt" <tedm@xxxxxxxxxxxxxxxxxxx>
 - *Date:* Mon, 16 Oct 2006 21:44:11 -0700
-

----- Original Message -----

From: "Erik Norgaard" <norgaard@xxxxxxxxxxxx>
To: "Ted Mittelstaedt" <tedm@xxxxxxxxxxxxxxxxxxx>
Cc: "Beech Rintoul" <freebsd@xxxxxxxxxxxxxxxxxxxx>;
<freebsd-questions@xxxxxxxxxxxx>
Sent: Sunday, October 15, 2006 3:47 AM
Subject: Re: Non English Spam

Ted Mittelstaedt wrote:

I have noted however, that some subscribers to this list write english encoded in one of the above character sets, I don't know enough about the character set definition, but it seems that English characters are

a

subset of any character set?

What is the recommended policy here? Should subscribers be advised to change character set when posting to the list?

No. It's the responsibility of the person doing the filtering – in this case you – to exempt any known good e-mail sender from your filters.

You know damn well that legitimate mailing list mail comes from mx2.freebsd.org (mx2.freebsd.org [216.136.204.119])

Re: Non English Spam

it's right in the headers of the messages on the list.

First: You know all too well that filtering based on "Received" header fields is not reliable – any decent spammer know how to forge that.

Spammers cannot forge the Received header that your own mailserver puts into the received message. The first Received line of the message is always legitimate. You can also turn on the Sendmail flag to put in the envelope address if you have multiple aliases to a mailbox that you want to see.

Accepting mail from a particular host should be done even before the mail delivery starts.

Don't know what your talking about here.

Second: If you know postfix, you also know that header filtering is independent of other checks, even the result of filtering on individual header lines are independent.

So the ideal you mention is not an option until a complete public list of authorized mail servers is available and all mail relayed through these requires authentication.

I don't know Postfix. So what your saying is Postfix is so defective that you can't use it for filtering? No wonder I never bothered to deal with it.

And, this isn't true anyway. You can easily tell with a little sleuthing what all of the mail emitters are for the FreeBSD mailing lists. Many mailing list managers, in fact, go to the trouble of posting publically what their mailservers are. And if the transmitting domain really has their shit together, they will have published SPF records in their DNS that will tell you what the authorized mailservers for that domain are. Sendmail has an SPF milter and I believe Spamassassin can also use these for weighting. (I'm too lazy to check for sure right now)

Or do you have the solution that does not imply accepting any of a myriad of character sets?

I'd be happy to implement that, but I don't want to open my mail server

Re: Non English Spam

to receive mail I have no means of reading and understanding just because it is RFC compliant.

You open your mailserver to known, whitelisted, legitimate sending servers, and let everyone else deal with the charset filtering. You know your going to accept mail from freebsd.org (or you tell your users to tell you if they are) and you exempt these servers from filtering.

You have no right to force other people to conform to what you feel is acceptable formatting of their message as long as they meet the SMTP rfc standards. That's why we have RFC's.

You you know perfectly well that content filtering is not based on the RFC's on SMTP but rather on the Internet Message Format and various RFC's on MIME – but I assume that you meant to refer to these.

content filtering and message charsets aren't the same thing. A content filter checks for "Make Money FAST" and other obvious spam content. You don't like Viagra? That's a content filter that takes care of that.

However, marking a non–spam message as spam soley because it's written in another language that you don't read – that's not a content filter. There's nothing in the content of that message that is spam. Thus you have no moral right to force mailing list users to conform to a specific language. Certainly, you can say "I don't know Spanish so I will just setup a filter to delete anything I get written in Spanish" but your crossing the line when you start telling people they can't post Spanish to a mailing list.

Basically what you say here is that spammers have every right to flood mail servers as long as they do so compliant with the RFC's?

I'm saying that you don't have the right to force other people to modify their content on messages that AREN'T spam just because your spam filters are too piss–poor to differentiate between an Asian charset message that is spam, and an Asian charset message that is a legitimate message.

Re: Non English Spam

I don't force anyone to conform to any arbitrary standards that I decide upon, but I have every legitimate right to reject anything that doesn't conform to my arbitrary standards.

No argument there – but your crossing the line (or the other poster is crossing the line) when your talking about telling list subscribers to change charsets when they post.

Keep in mind also that the entire point of the list is for newbies anyway, and there's been plenty of busted mail client software out there in the past that mishandled charsets even on English, and there will be in the future. Some of these people wouldn't know a charset from a chair, much less how to change it, and are coming from Windows systems and God-knows what else. What do you want to do with these folks? Ignore them so they think we are a bunch of arseholes and then go to the Penguin? Or guide them in the paths of righteousness?

Yet, it is somewhat implicit that this is an English language list, any one writing in a different language may be lucky to find someone who can respond in their language, but are just as often referred to one of the language specific lists – if their message is not simply ignored.

Yes, that is true. Thus, the best way to handle non-English posts to the list by non-English posters is to reply, off-list, and direct the subscriber elsewhere, or tell them to re-post in English.

That is how just about all mailing lists have handled this in the past and how they all handle it now, and there is no good reason to start setting up filters on the mailing list to block these kinds of posts.

So we do actually impose some arbitrary rule on subscribers, namely to write in English.

No, we don't. If non-English speakers were to carry on a dialog on the mailing list it would be perfectly fine. It has, in fact, happened in the past that non-English speakers have posted here and had responses in their native language that answered their question, along with the usual statement that they would get a lot more assistance if they posted their questions in English. It's not that common, but it has happened a few times that I've seen.

Given that we find it reasonable to impose such a rule,

Re: Non English Spam

That is the heart of the debate. You think it's reasonable simply because you decided Sendmail was too difficult to learn and copped out on an easy replacement – that replacement is now falling down on the increased complex demands spammers present, and your going to try to make the world conform to what the crap software your using can handle.

It isn't a given that this is a reasonable rule to impose and nothing you have said so far supports the assertion that it is.

then why is it unreasonable to impose that they should abstain from obscure non-English character sets?

I was hoping to find a way that we can all get along, I find it kind of useless to waste my resources on mail written in languages that I have no means of interpreting.

But, since you already said that your filtering non-English postings, your basically being completely silly since the percentage of legitimate non-English mails you get is probably a ten thousandth of the spam – having the FreeBSD mailserver drop it before you get it won't make a dent in the resources your already spending to filter.

If everyone did what your proposing then senders would have hundreds of different rules they would have to follow, over and above the normal RFCs.

Well, in real life as well as on-line we have thousands of rules and customs, implicit or written, on communication and gestures.

Oh baloney. Some of the biggest flame wars I've seen started because people LACK the many thousands of non-verbal feedback cues that are present in standard verbal or face to face communication.

Email is significant in what it LACKS in the way of rules or customs, not what it HAS.

There are best practices on how to communicate in e-mail and on mailing lists, usage of smileys and other types of mood-expression, and proclaimed best practices on how to quote.

Re: Non English Spam

There's a few people out there who claim to tell everyone else how to send e-mails, yes. But you can turn on the TV and see the Dr. Phil show and see the same thing for face to face communication. Both are for amusement value only.

You regularly see people complaining about top posting. Then, line wrapping, or people who don't delete the trailing message part that they don't reply to etc.

No, not at all. What I DO regularly see is people who don't like the CONTENT of a message – they don't like being called a stupid arsehole when they are being one, for example – and because they have no real defense for acting the stupid arsehole, they instead retaliate by complaining about top posting, or spelling, or some other lamers whine.

Nobody who ever really LIKED the content of a message EVER bitched out the poster about top posting. Oh sure, they will complain if they don't like the content – but if the responder is helping THEM with THEIR little problem, well then by God, top post away, all day long and more power to you!!!

Ted

freebsd-questions@xxxxxxxxxxx mailing list

<http://lists.freebsd.org/mailman/listinfo/freebsd-questions>

To unsubscribe, send any mail to "freebsd-questions-unsubscribe@xxxxxxxxxxx"