

Talkback: Reclaiming the Blogosphere

Elie Bursztein, Baptiste Gourdin, John Mitchell
Stanford University & LSV-ENS Cachan

忍
友
喜
愛
幸
福

福
愛

友
喜
愛
幸
福

What is a blog ?

- A Blog ("Web log") is a site, usually maintained by an individual with
 - Regular entries
 - Commentary
 - **LinkBack**
- Entries displayed in reverse-chronological order.



The screenshot shows a blog page with a dark header containing the text "from information to intelligence" and a navigation menu with links for "Blog", "Publications", "Talks", "Contact", and "Security And Privacy Highlights". The main content area features two articles. The first article, titled "Display your latest Foursquare checkins on your blog using Foursquare API v2 and PHP", includes a Foursquare logo and a "Read More..." link. The second article, titled "Using the Microsoft Geolocation API to retrace where a Windows laptop has been", includes a map of North Las Vegas and an "EDIT" section dated Tuesday 2nd August. The sidebar on the right contains a profile for Elie Bursztein, a search bar, social media icons for RSS, Twitter, Facebook, and Google+, and sections for "LATEST WEEK HIGHLIGHTS" and "POPULAR POSTS".

<http://elie.im/blog>

Key Statistics

- 184 Millions blogs
- 73% of users read blogs
- 50% post comments



universalmccann

Anatomy of a blog post

The image shows a screenshot of a WordPress blog post editor. The post title is "New MailChimp Logo By Jon Hicks". The permalink is "http://www.mailchimp.com/blog/new-mailchimp-...o-by-jon-hicks/". The post content includes a paragraph about redesigning the MailChimp logo and a sketch of a character. The categories are "IMHO" and "MailChimp News". The excerpt is "The story behind MailChimp's new branding, including sketches from Jon Hicks". The post author is "Ben".

1. Title

2. Permalink

3. Body / Content

4. Publish Date

5. Category

6. Summary / Excerpt

7. Author

View this Post

Publish Status

Published

Published on:
September 11, 2008 at 4:49 am [Edit](#)

Save [Delete post](#)

Last edited by Ben on September 11, 2008 at 4:49 am
Draft Saved at 10:01:29 pm.
Word count: 1492

Related

- See Comments on this Post
- Manage All Comments
- Manage All Posts
- Manage All Categories
- Manage All Tags
- View Drafts

Shortcuts

Drag-and-drop the following link to your bookmarks bar or right click it and add it to your favorites for a posting shortcut. [Press This](#)

Advanced Options

Excerpt

The story behind MailChimp's new branding, including sketches from Jon Hicks

Excerpts are optional hand-crafted summaries of your content. You can [use them in your template](#)

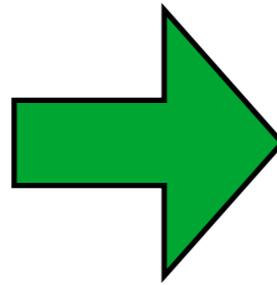
Post Author

Ben

Why blogs are special ?



User



Elie Bursztein Web, mobile and offensive security research

Home About me Publications Talks Software Blog

Social news

- 29 minutes ago RT @jeremiahg: "Expression Language Injection" ..."new, specific attacks against Spring MVC JSP tags" <http://t.co/unx26Vn> #security #java
- about 3 hours ago Checking out <http://t.co/kppE3NQ> My summary of last week security and privacy top headlines is out #security #privacy
- 3 days ago Checking out <http://t.co/61MLajV> Firesheep addon updated to exploit Google info leak: 40% of your click history can be stolen. #privacy #security #www
- 3 days ago Reading of the day: <http://t.co/SzuUGsV> How Researchers used Typosquatting to stole 20 GB of E-Mail From Fortune 500 #security #privacy
- 3 days ago Reply to @myonline Sorry :) My ipad new reader tends to remove the HTTP prefix sometime for no reason ... #EPICfail

Blog Posts

- about 4 hours ago Security and privacy top headlines September 4th ? 10th
- 6 days ago Security and privacy top headlines
- 17 days ago ...

Selected recent publications

- Towards Secure Embedded Web Interfaces**
by Baptiste Gourdin Chinmay Soman Hristo Bojinov Elie Bursztein
Usenix Security 2011, San Francisco, USA
Tags: web, embedded devices, offensive technologies
- OpenConflict: Preventing Real Time Map Hacks in Online Games**
by Elie Bursztein Jocelyn Lagarenne Mike Hamburg Dan Boneh
S&P 2011, Oakland, USA
Tags: game, crypto, offensive technology
- The Failure of Noise-Based Non-Continuous Audio Captchas**
by Elie Bursztein Romain Bauxis Hristo Paskov Daniele Perito Celine Fabry John C. Mitchell
S P 2011, Oakland, USA
Tags: captcha, web, offensive technology

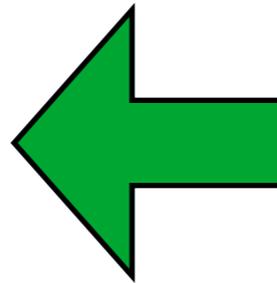
Selected recent talks

- Beyond files recovery, OWADE cloud-based forensic**
BlackHaT USA 2011, Las Vegas, USA
Tags: forensic, windows
- Kartograph**
Defcon 18, Las Vegas, USA
Tags: game, offensive technologies

Why blogs are special ?



User



The screenshot shows a blog page with a dark header containing the title 'From information to intelligence' and navigation links: 'Blog', 'Publications', 'Talks', 'Contact', and 'Security And Privacy Highlights'. The main content area features two articles. The first article is titled 'Display your latest Foursquare checkins on your blog using Foursquare API v2 and PHP', dated 'Aug, 26 2011', with '8 comments'. It includes a blue and purple Foursquare logo icon. The second article is titled 'Using the Microsoft Geolocation API to retrace where a Windows laptop has been', dated 'Jul, 29 2011', with '51 comments'. It includes a map icon. To the right of the articles is a sidebar with a profile picture of Elie Bursztein, a search bar, social media icons for RSS, Twitter, Facebook, and Google+, and sections for 'LATEST WEEK HIGHLIGHTS' and 'POPULAR POSTS'.

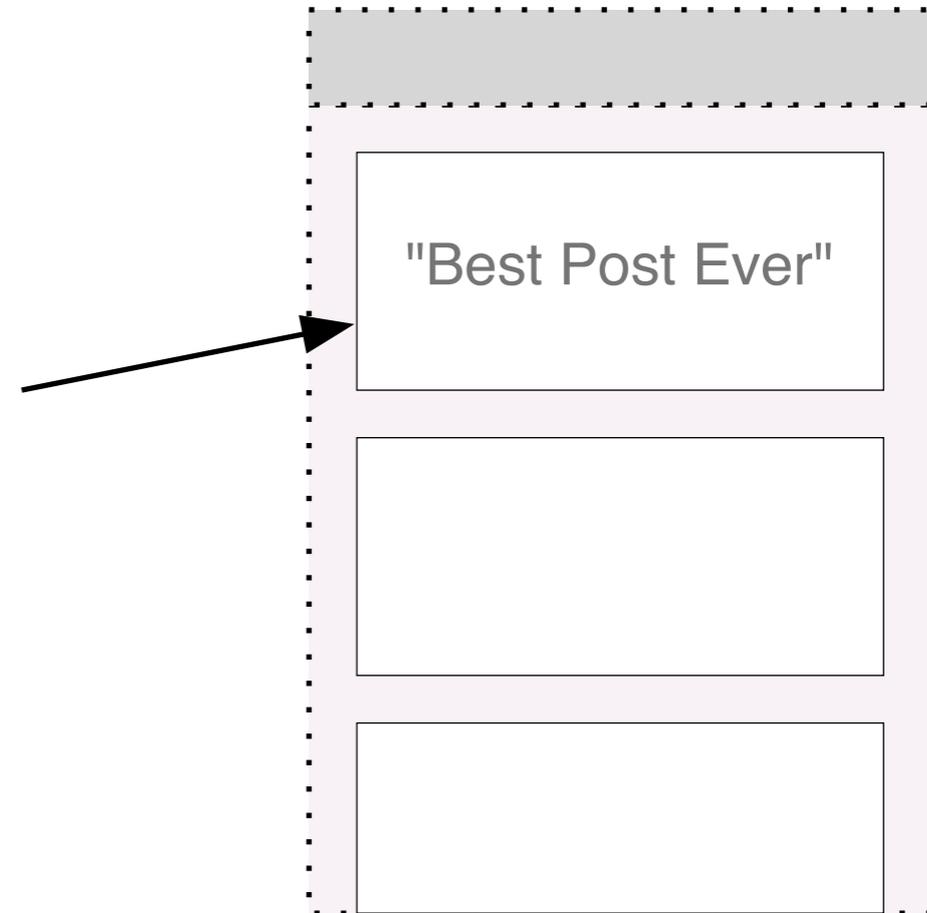
What is a TrackBack ?



Trackback Illustrated

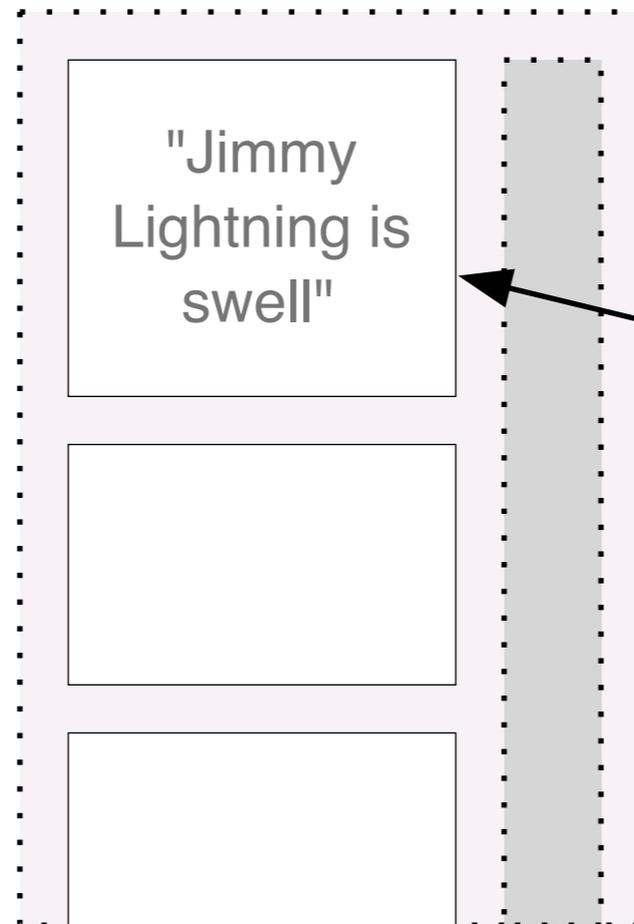
1. ■

"Wow! Jimmy Lightning has written the best post ever! It's so funny! And it's true! That's why it's so good. I need to tell the world!"



Trackback Illustrated

2.



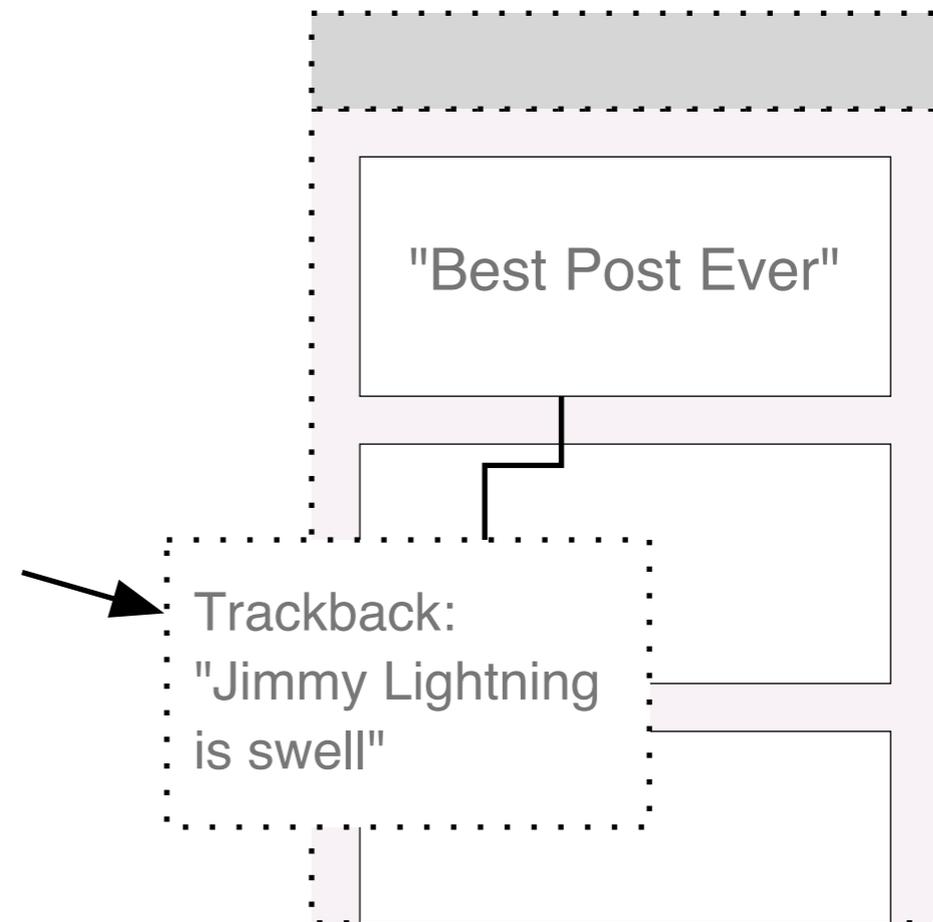
"Check it out world! I've written all about Jimmy Lightning's post on my weblog. My weblog's called '*The Unbloggable Blogness of Blogging*'. It's a good name huh? Wow, I sure hope Jimmy sees what I said about him..."

Trackback Illustrated

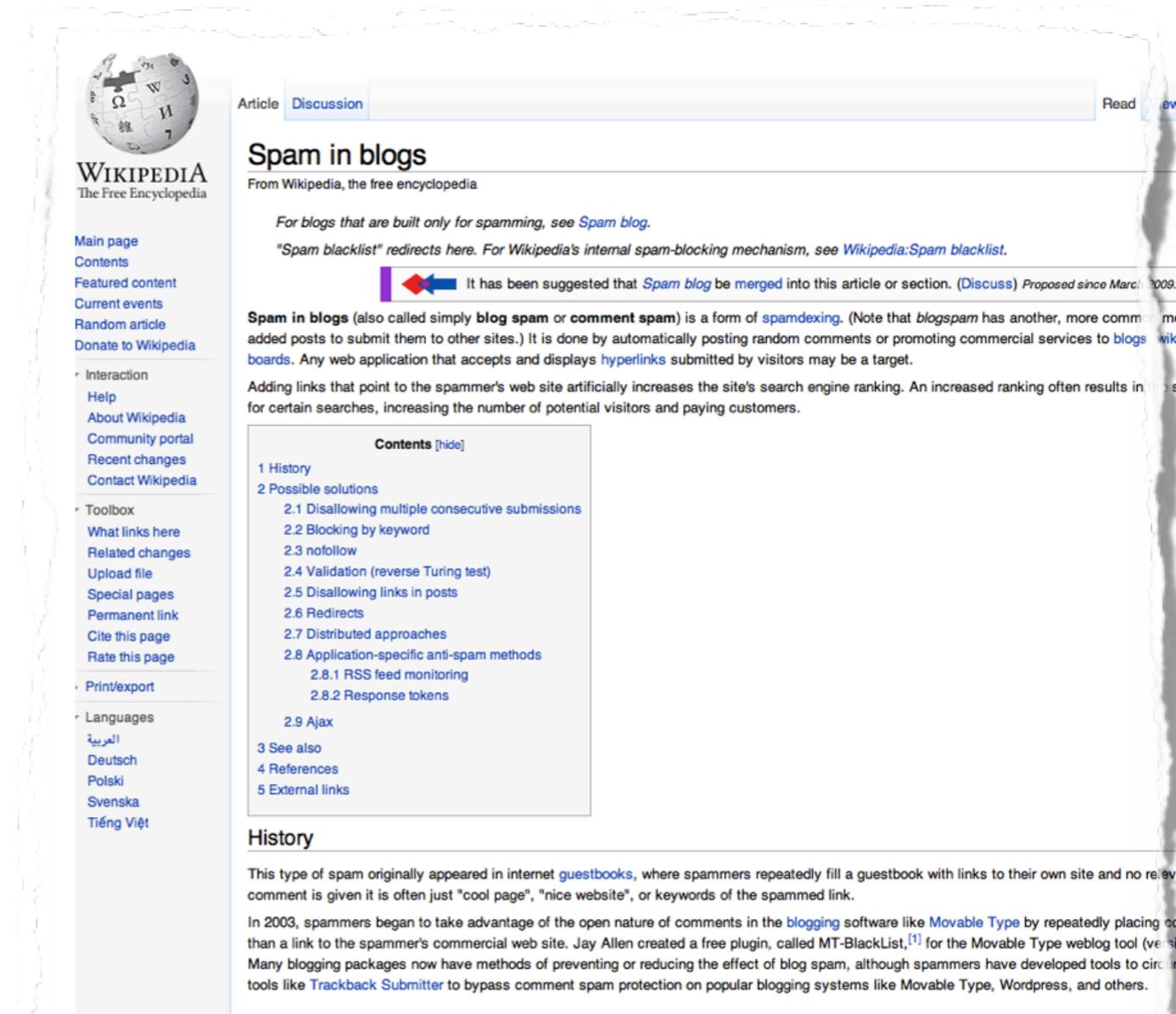
3.

"Hey! I wonder what happens when I click on that Trackback link at the bottom of Jimmy's post?"

"Far out! In that funny pop-up is a link back to my site! Too cool for school!"



Why the LinkBack problem is different ?



The screenshot shows a Wikipedia article page for "Spam in blogs". The page layout includes a left sidebar with navigation links, a main content area with a title, a sub-header, and a table of contents. The article text discusses the nature of spam in blogs and its impact on search engine rankings.

WIKIPEDIA
The Free Encyclopedia

Article [Discussion](#) Read [new](#)

Spam in blogs

From Wikipedia, the free encyclopedia

For blogs that are built only for spamming, see [Spam blog](#).

"Spam blacklist" redirects here. For Wikipedia's internal spam-blocking mechanism, see [Wikipedia:Spam blacklist](#).

 It has been suggested that [Spam blog](#) be merged into this article or section. ([Discuss](#)) *Proposed since March 2009.*

Spam in blogs (also called simply **blog spam** or **comment spam**) is a form of [spamdexing](#). (Note that *blogspam* has another, more common meaning: [spamming](#) [wiki](#) [boards](#). Any web application that accepts and displays [hyperlinks](#) submitted by visitors may be a target.

Adding links that point to the spammer's web site artificially increases the site's search engine ranking. An increased ranking often results in more visitors to the site, increasing the number of potential visitors and paying customers.

Contents [hide]

- 1 History
- 2 Possible solutions
 - 2.1 Disallowing multiple consecutive submissions
 - 2.2 Blocking by keyword
 - 2.3 nofollow
 - 2.4 Validation (reverse Turing test)
 - 2.5 Disallowing links in posts
 - 2.6 Redirects
 - 2.7 Distributed approaches
 - 2.8 Application-specific anti-spam methods
 - 2.8.1 RSS feed monitoring
 - 2.8.2 Response tokens
 - 2.9 Ajax
- 3 See also
- 4 References
- 5 External links

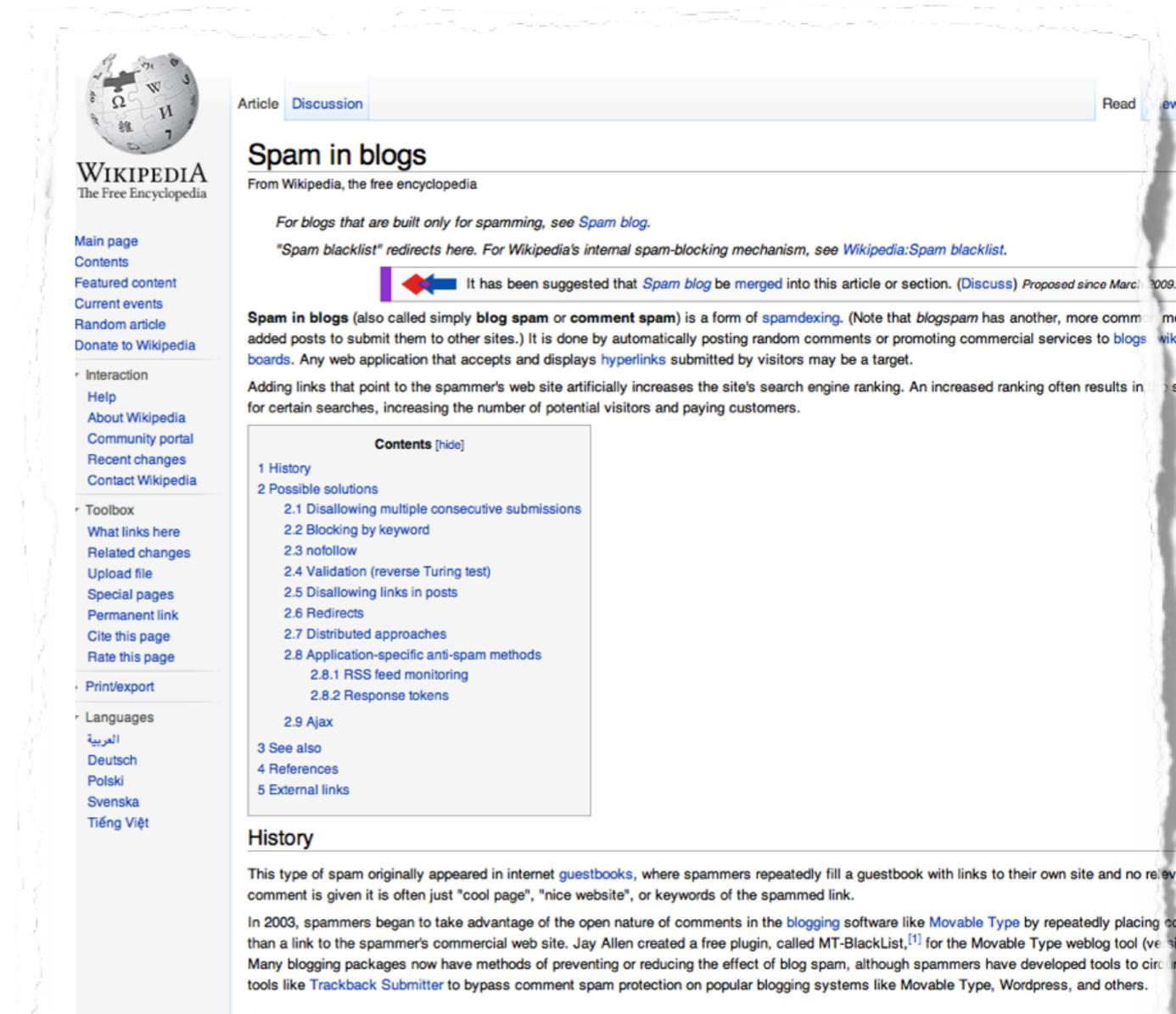
History

This type of spam originally appeared in internet [guestbooks](#), where spammers repeatedly fill a guestbook with links to their own site and no relevant comment is given it is often just "cool page", "nice website", or keywords of the spammed link.

In 2003, spammers began to take advantage of the open nature of comments in the [blogging](#) software like [Movable Type](#) by repeatedly placing comments containing more than a link to the spammer's commercial web site. Jay Allen created a free plugin, called [MT-BlackList](#),^[1] for the Movable Type weblog tool (version 3.0). Many blogging packages now have methods of preventing or reducing the effect of blog spam, although spammers have developed tools to circumvent these tools like [Trackback Submitter](#) to bypass comment spam protection on popular blogging systems like Movable Type, Wordpress, and others.

Why the LinkBack problem is different ?

- A **single** spam can reach **thousand** of users



The screenshot shows a Wikipedia article page for "Spam in blogs". The page includes a sidebar with navigation links, a main content area with a table of contents, and a history section. The table of contents lists sections such as "History", "Possible solutions", "See also", "References", and "External links". The history section contains text explaining the origin of spam in internet guestbooks and the development of anti-spam measures in blogging software.

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox
What links here
Related changes
Upload file
Special pages
Permanent link
Cite this page
Rate this page

Print/export

Languages
العربية
Deutsch
Polski
Svenska
Tiếng Việt

Article Discussion

Spam in blogs

From Wikipedia, the free encyclopedia

For blogs that are built only for spamming, see Spam blog.

"Spam blacklist" redirects here. For Wikipedia's internal spam-blocking mechanism, see Wikipedia:Spam blacklist.

It has been suggested that *Spam blog* be merged into this article or section. (Discuss) Proposed since March 2009.

Spam in blogs (also called simply **blog spam** or **comment spam**) is a form of *spamdexing*. (Note that *blogspam* has another, more common meaning: a collection of spamme

added posts to submit them to other sites.) It is done by automatically posting random comments or promoting commercial services to **blogs** and **wiki** boards. Any web application that accepts and displays **hyperlinks** submitted by visitors may be a target.

Adding links that point to the spammer's web site artificially increases the site's search engine ranking. An increased ranking often results in more visitors to the site, increasing the number of potential visitors and paying customers.

Contents [hide]

- History
- Possible solutions
 - Disallowing multiple consecutive submissions
 - Blocking by keyword
 - nofollow
 - Validation (reverse Turing test)
 - Disallowing links in posts
 - Redirects
 - Distributed approaches
 - Application-specific anti-spam methods
 - RSS feed monitoring
 - Response tokens
 - Ajax
- See also
- References
- External links

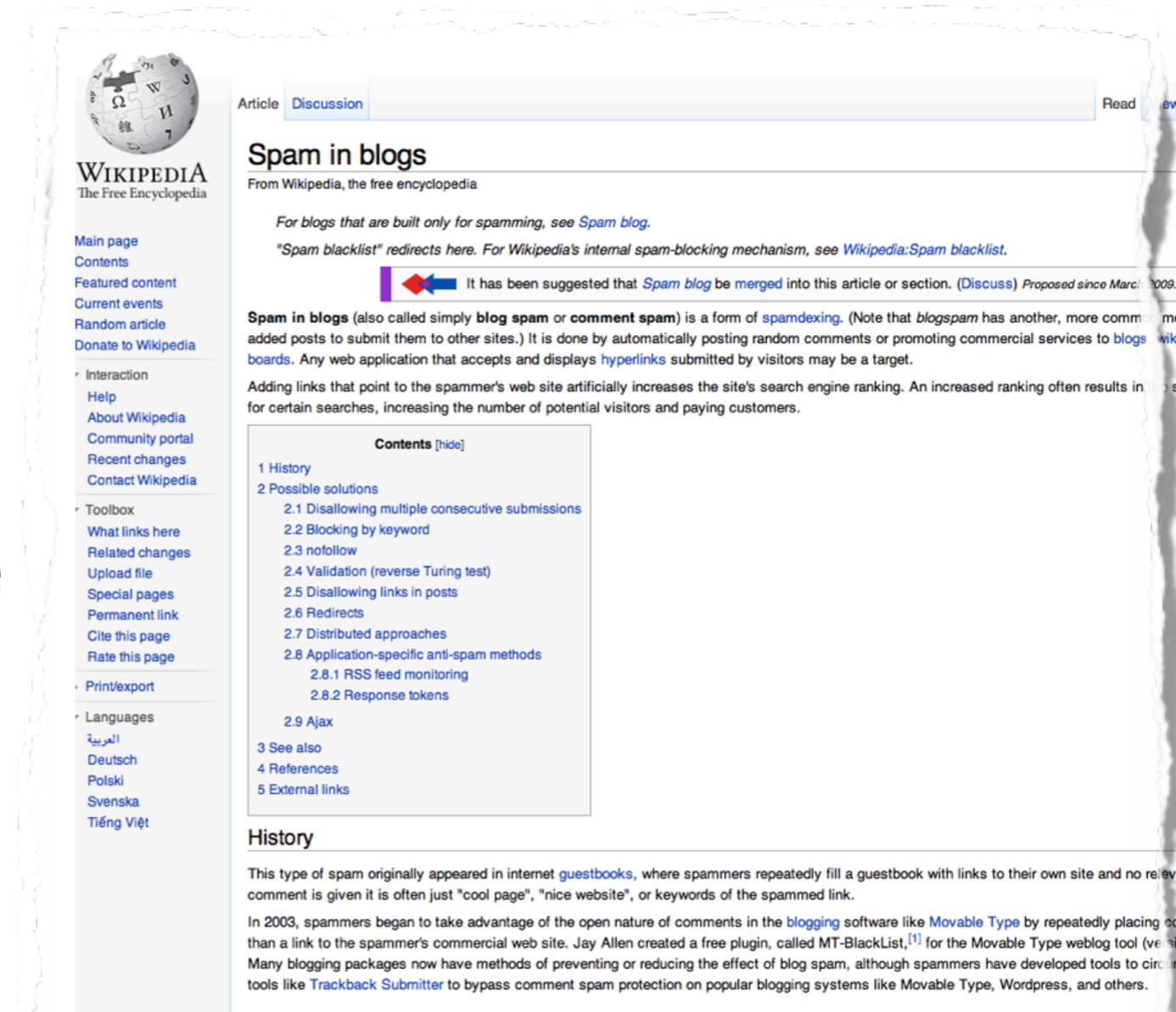
History

This type of spam originally appeared in internet *guestbooks*, where spammers repeatedly fill a guestbook with links to their own site and no relevant comment is given it is often just "cool page", "nice website", or keywords of the spammed link.

In 2003, spammers began to take advantage of the open nature of comments in the *blogging* software like *Movable Type* by repeatedly placing comments containing more than a link to the spammer's commercial web site. Jay Allen created a free plugin, called MT-BlackList,^[1] for the Movable Type weblog tool (version 3.0). Many blogging packages now have methods of preventing or reducing the effect of blog spam, although spammers have developed tools to circumvent these tools like *Trackback Submitter* to bypass comment spam protection on popular blogging systems like Movable Type, Wordpress, and others.

Why the LinkBack problem is different ?

- A **single** spam can reach **thousand** of users
- LinkBack **notification** are **automated**



The screenshot shows a Wikipedia article page for "Spam in blogs". The page includes a sidebar with navigation links, a main content area with a table of contents, and a history section. The table of contents lists sections such as "History", "Possible solutions", "See also", "References", and "External links". The history section explains that this type of spam originally appeared in internet guestbooks and that spammers began to take advantage of the open nature of comments in blogging software like Movable Type in 2003.

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox
What links here
Related changes
Upload file
Special pages
Permanent link
Cite this page
Rate this page

Print/export

Languages
العربية
Deutsch
Polski
Svenska
Tiếng Việt

Article Discussion

Spam in blogs

From Wikipedia, the free encyclopedia

For blogs that are built only for spamming, see Spam blog.

"Spam blacklist" redirects here. For Wikipedia's internal spam-blocking mechanism, see Wikipedia:Spam blacklist.

It has been suggested that *Spam blog* be merged into this article or section. (Discuss) Proposed since March 2009.

Spam in blogs (also called simply **blog spam** or **comment spam**) is a form of *spamdexing*. (Note that *blogspam* has another, more common meaning: a large number of unwanted or irrelevant posts to submit them to other sites.) It is done by automatically posting random comments or promoting commercial services to *blogs* or *wiki boards*. Any web application that accepts and displays *hyperlinks* submitted by visitors may be a target.

Adding links that point to the spammer's web site artificially increases the site's search engine ranking. An increased ranking often results in more traffic to the site, increasing the number of potential visitors and paying customers.

Contents [hide]

- History
- Possible solutions
 - Disallowing multiple consecutive submissions
 - Blocking by keyword
 - nofollow
 - Validation (reverse Turing test)
 - Disallowing links in posts
 - Redirects
 - Distributed approaches
 - Application-specific anti-spam methods
 - RSS feed monitoring
 - Response tokens
 - Ajax
- See also
- References
- External links

History

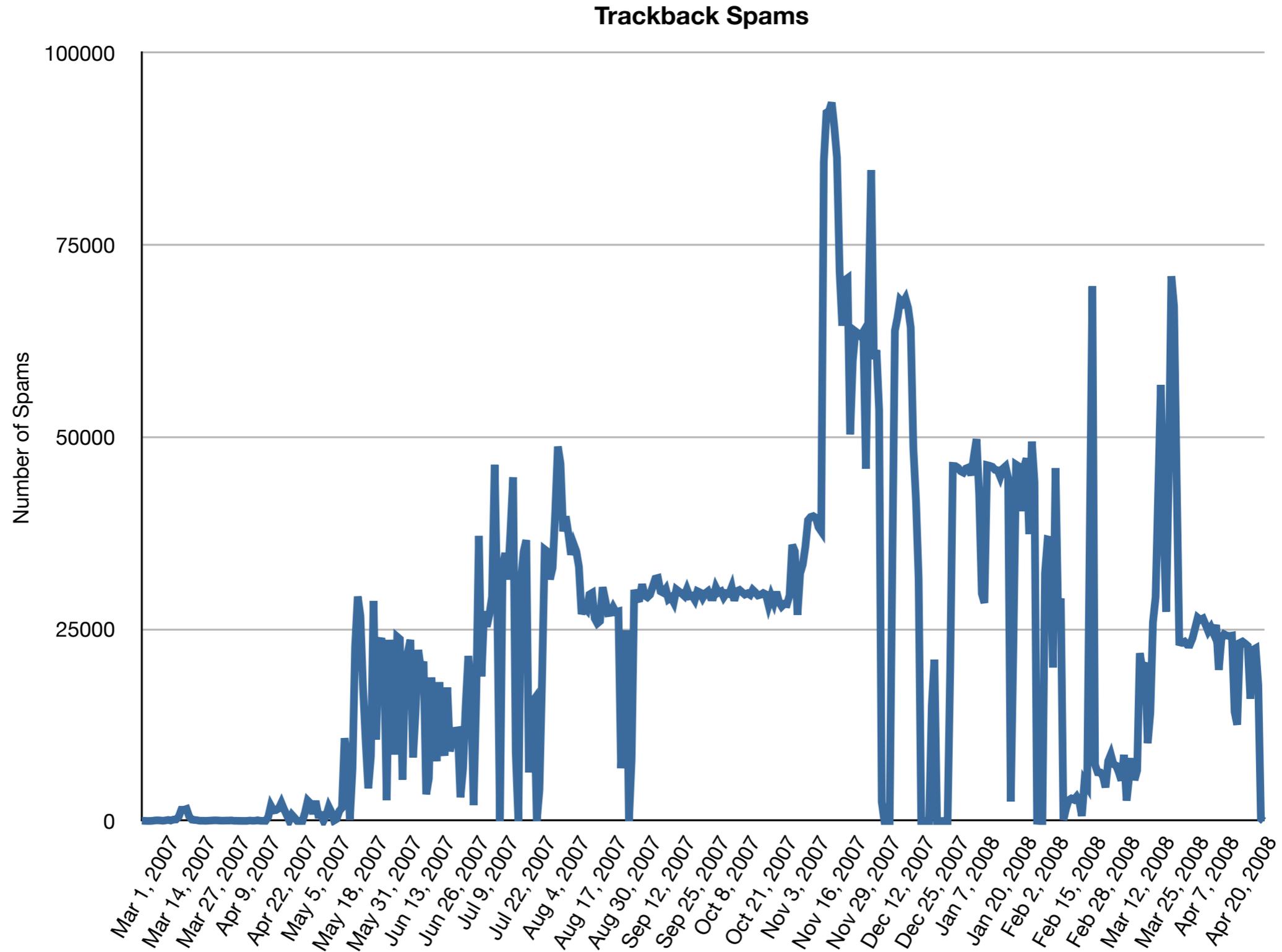
This type of spam originally appeared in internet *guestbooks*, where spammers repeatedly fill a guestbook with links to their own site and no relevant comment is given it is often just "cool page", "nice website", or keywords of the spammed link.

In 2003, spammers began to take advantage of the open nature of comments in the *blogging* software like *Movable Type* by repeatedly placing comments containing more than a link to the spammer's commercial web site. Jay Allen created a free plugin, called MT-BlackList,^[1] for the Movable Type weblog tool (version 3.0). Many blogging packages now have methods of preventing or reducing the effect of blog spam, although spammers have developed tools to circumvent these tools like *Trackback Submitter* to bypass comment spam protection on popular blogging systems like Movable Type, Wordpress, and others.



30 million LinkBack spam by **day**

Spam Campaign example



The end of all hope ?



CODING HORROR

programming and human factors
by Jeff Atwood

Dec 20, 2006

The Day The Trackbacks Died

You might read a post on this blog and decide I feel this way. And, while you're at it, feel free to public dialog is why I believe so strongly in com

But sometimes a mere comment isn't enough. I write an entire post on *your* blog explaining, in you know that I, the target of your vitriol, have a little too intimate. Or, you could leave a com

NEWER »

[Assertiveness for Software Developers](#)

OLDER

[Eric Lippert's Purple Crayon](#)

[Home](#)

[Browse All Posts](#)

The end of all hope ?



CODING HORROR

programming and human factors
by Jeff Atwood



NEWER »

[Assertiveness for Software Developers](#)

OLDER

[Eric Lippert's Purple Crayon](#)

[Home](#)

[Browse All Posts](#)

inn covers a bit of everything tech with a dash of environment and world news. It was written by Patrick Tanguay, a professional web developer and consultant in Montréal who also blogs pictures and opinions at [Céboça](#) and is editor in chief of [NextMontréal](#). He works out of [Station](#) a working space he co-founded.

Follow me on Twitter; [@inevernu](#)

[Subscribe to RSS](#) [Tags](#)

ES [Subscribe](#)

Trackback's Dead, Take It Off Your Site

We've known for a while that trackback is broken. Initially it was an annoyance, but then, instead of writing a long winded (perhaps somewhat off topic) comment, you write it on your own blog and "ping" the original post. Then people started threading the comments and trackbacks with the comment display block instead of separate. Then spam arrived.

Now not only have a lot of people stopped offering trackback urls on their blogs, but most have stopped using the system completely, preferring the simpler and more effective now that referral tools and services like [Pubsub](#) and [Technorati](#) are making it easy to find who's linking you.

So, having said all of that, when's the last time you've seen a trackback brought anything at all to a discussion? Actually, when's the last time you've clicked on a trackback link, except maybe by mistake, thinking it was a comment and did it this morning for the first time in maybe as much as a year. Not

The end of all hope ?



CODING HORROR

programming and human factors
by Jeff Atwood

NEWER »

[Assertiveness for Software Developers](#)

OLDER

[Eric Lippert's Purple Crayon](#)

[Home](#)

[Browse All Posts](#)



inn covers a bit of everything tech with a dash of environment and world news. It is written by Patrick Tanguay, a professional web developer and consultant in Montréal who also blogs pictures and stories at [Céboça](#) and is editor in chief at [NextMontréal](#). He works out of [Station](#) working space he co-founded.

Follow inn on Twitter; [@inevernu](#)

[Subscribe to RSS](#) [Tags](#)

Be the first to know about inn

the BLOG HERALD

[news](#) • [editorial](#) • [features](#) • [interviews](#) • [guides](#) • [reviews](#) • [100 stories](#) • [archive](#) • [books](#) • [advertise](#)

Have Trackbacks Become Too Spammy To Be Worthwhile?

You're at: [Home](#) • [Editorial](#) • [Have Trackbacks Become Too Spammy To Be Worthwhile?](#) • Filed as [Editorial](#), [Features](#) on November 24, 2007 8:56 am

by **Tony Hung**



0



0

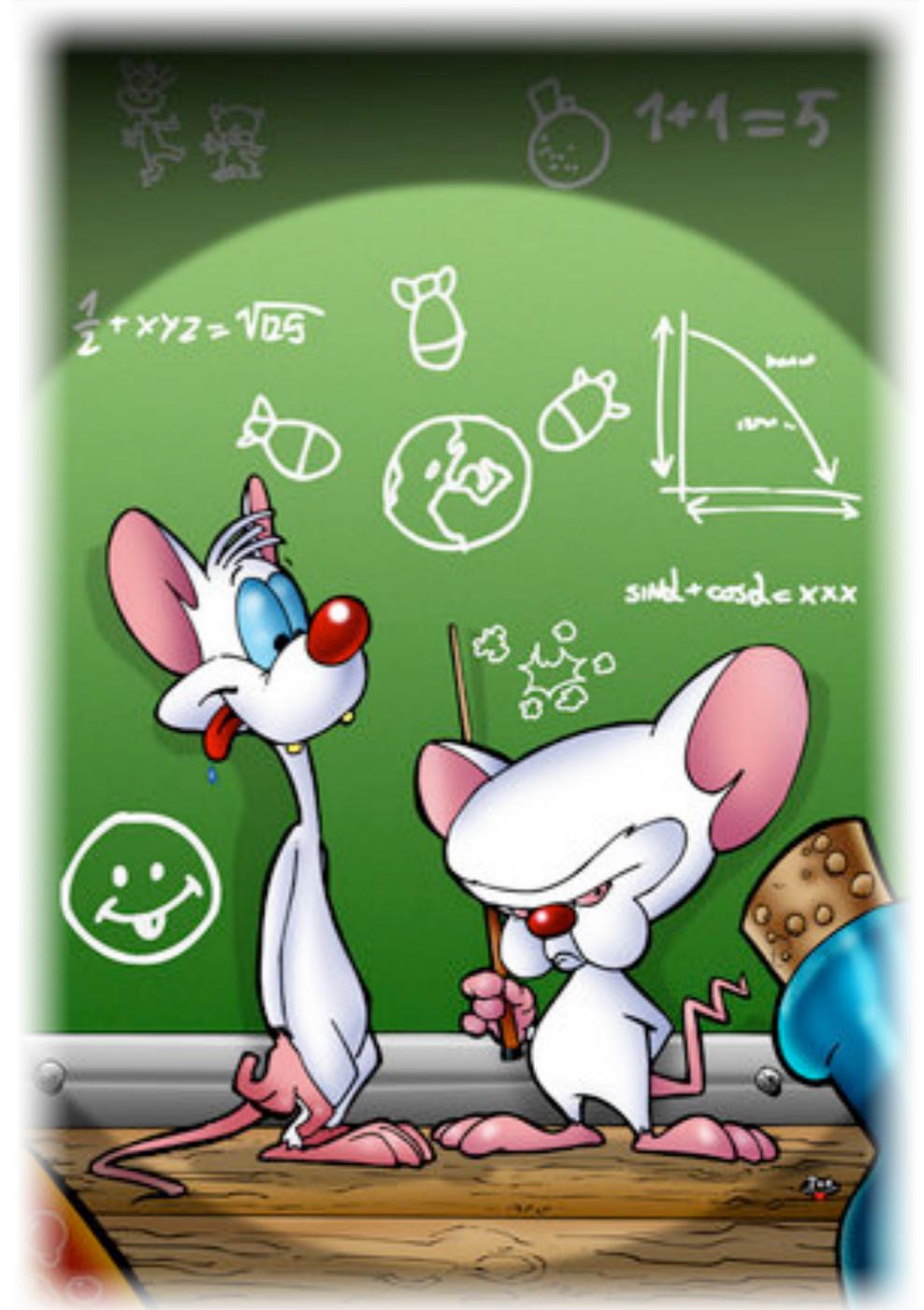


Be the first of your friends to like this.

You may have also discovered a surge in trackback spam recently as autoblogging software is being used by more and more spammers to reach out and cull RSS feeds. This phenomenon has led to many disabling trackbacks, or raising the “blacklist” level so high that you might never see some trackbacks again. Or, as some newer remotely-hosted commenting technologies like [IntenseDebate](#) and [Disqus](#) show, they simply do not show trackbacks because of the spam problem.

Adversary Model

- Resourceful
- Efficient
- Knowledgable
- Adaptive



Threat

- Blog spoofing
- Cried Wolf attack
- Linkback tampering and replay
- Spam in breath or in deapth



Current Linkback specifications

	RefBack	PingBack	TrackBack
Trigger Mechanism	Visit from the sender site	Code executed at posting time	Code executed at posting time
Notification medium	HTTP referer	XML-RPC call	HTTP POST
Information sent	none	<ul style="list-style-type: none">- <i>S</i> post URL- <i>R</i> post URL	<ul style="list-style-type: none">- <i>S</i> post URL- <i>S</i> site name- <i>S</i> post title- <i>S</i> post excerpt
Auto-discovery mechanism	none	LINK Tag	Specially formatted information in the body



Threats are **not addressed** by current Linkback specifications

The background of the slide is a close-up of red, vertically pleated curtains. The top edge shows the curtains gathered into a decorative valance. A dark grey horizontal band is centered across the middle of the image, containing the text.

Introducing **Talkback** !

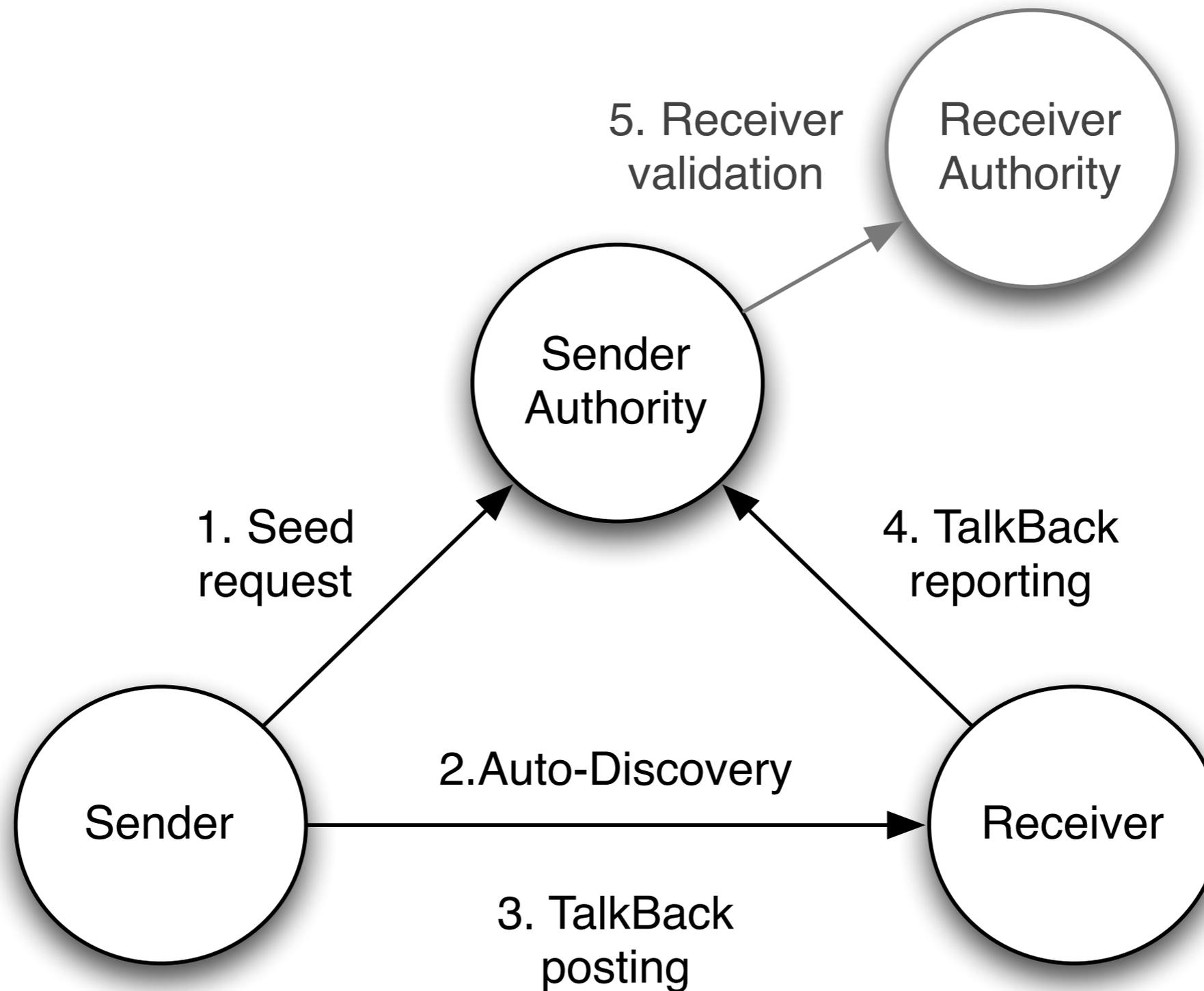


Use a **light-weight PKI** to fight
Spam

TalkBack Benefits

	RefBack	PingBack	TrackBack	TalkBack
Trigger Mechanism	Visit from the sender site	Code executed at posting time	Code executed at posting time	Code executed at posting time
Notification medium	HTTP referer	XML-RPC call	HTTP POST	HTTP POST
Information sent	none	<ul style="list-style-type: none"> - <i>S</i> post URL - <i>R</i> post URL 	<ul style="list-style-type: none"> - <i>S</i> post URL - <i>S</i> site name - <i>S</i> post title - <i>S</i> post excerpt 	<ul style="list-style-type: none"> - <i>S</i> post URL - <i>S</i> site name - <i>S</i> post title - <i>S</i> post excerpt - Seed Token - <i>S</i> Public key - <i>R</i> Public key - Signature
Auto-discovery mechanism	none	LINK Tag	Specially formatted information in the body	LINK Tag
<i>S</i> Authenticity	-	-	-	✓
<i>R</i> Authenticity	-	-	-	✓
Integrity	-	-	-	✓
Confidentiality	-	-	-	✓

Talkback Overview



Talkback Security Equation



Talkback Security Equation

Limited number of blogs

Talkback Security Equation

Limited number of blogs

+ Limited number of LinkBack by blog

Talkback Security Equation

Limited number of blogs

+ Limited number of LinkBack by blog

Spam under-control

Making blog registration costly

- Captcha
- Email verification
- Domain verification



LinkBack benefits

- **Blog**
 - Authenticity
 - Whitelisting / blacklisting
- **Linkback**
 - Integrity
 - Non-repudiable
 - Confidentiality (optional)

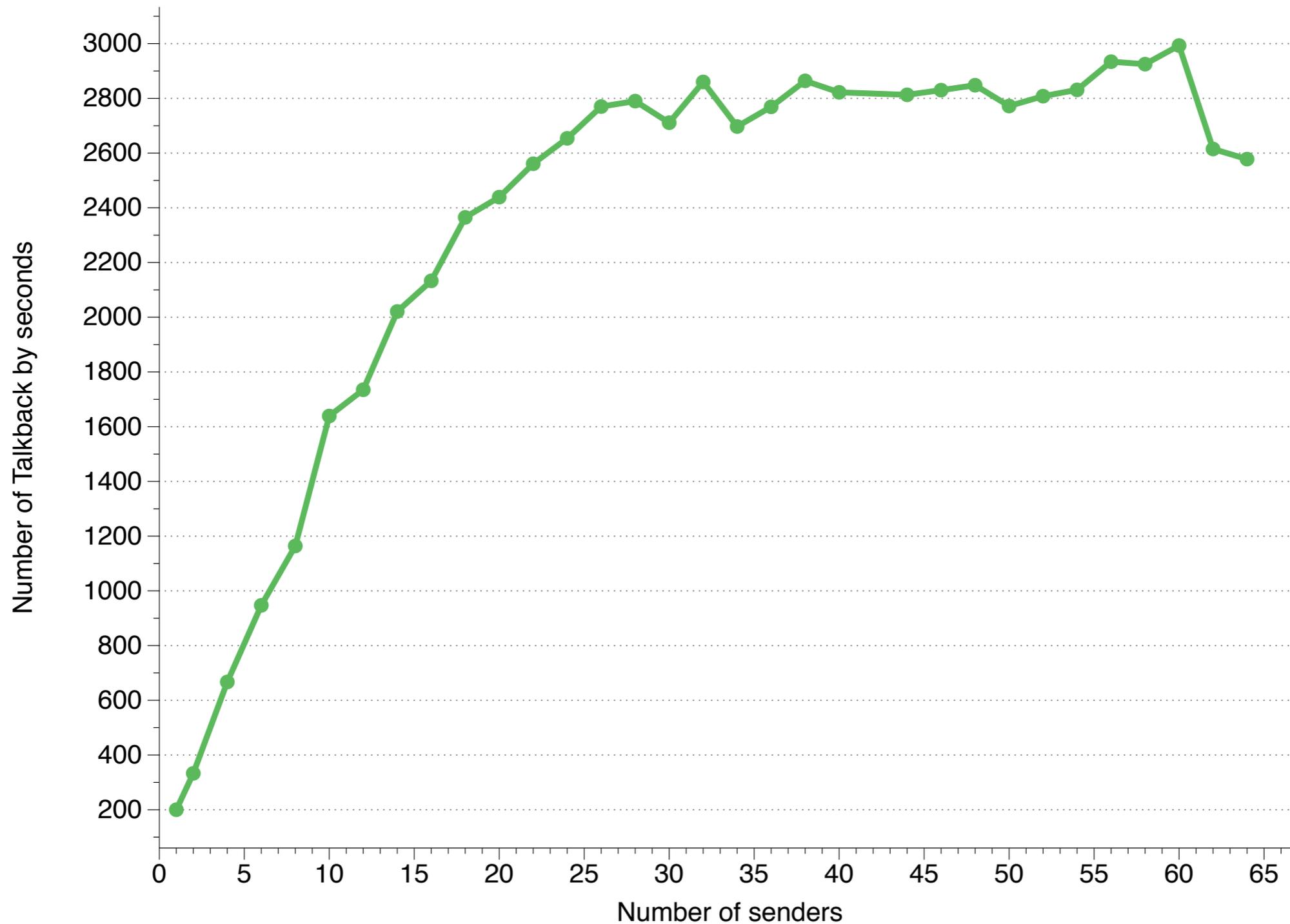
Reducing authority power

- Can't forge LinkBack: authorities have only blog public keys
- Respect privacy: authorities don't see LinkBack content in secure mode
- No single point of failure/control: protocol allows blog to choose the authority of their choice.

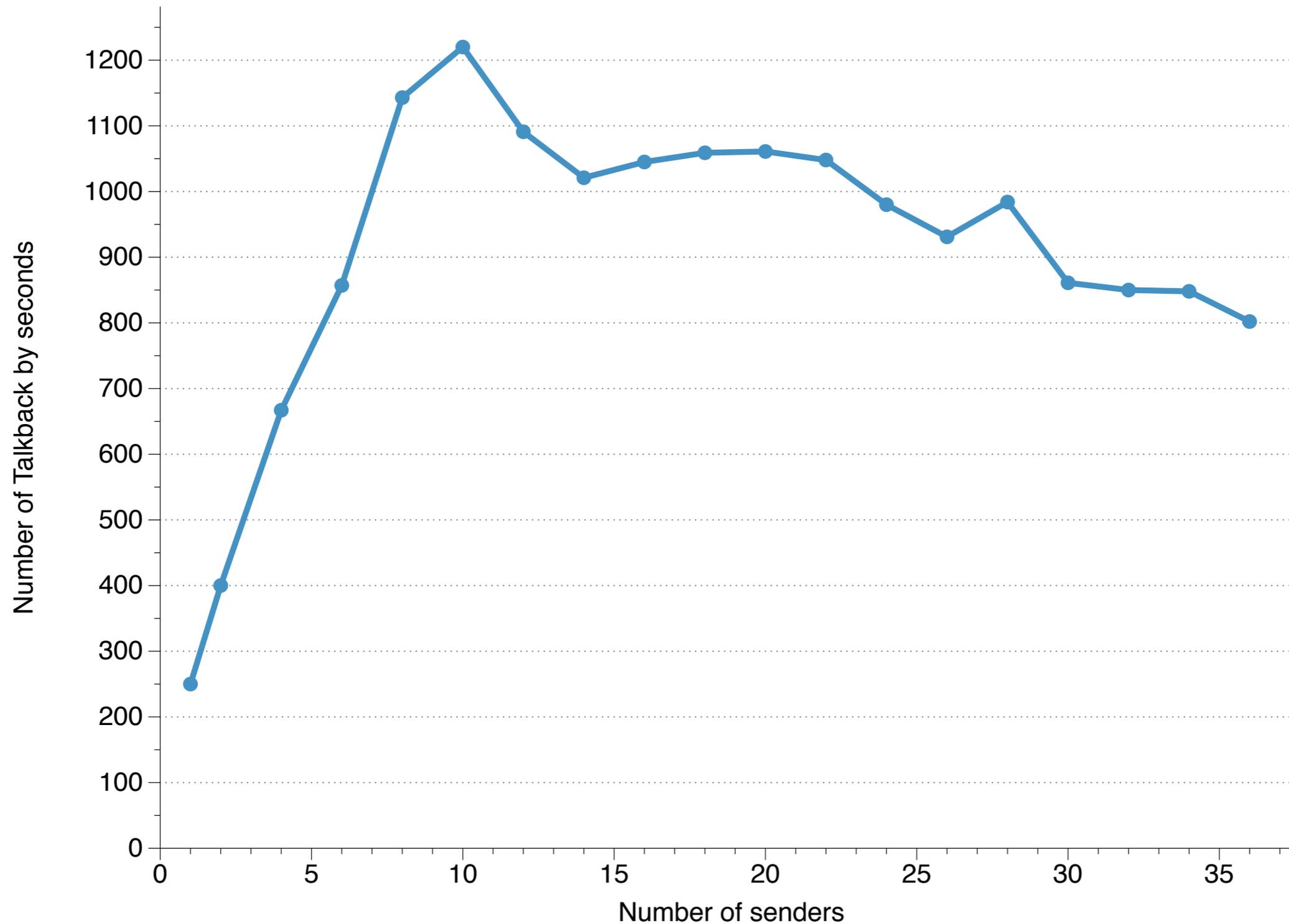
Implementation

- Open-source library
- Wordpress plugin
- Authority : <https://talkback.stanford.edu>

Authority benchmark



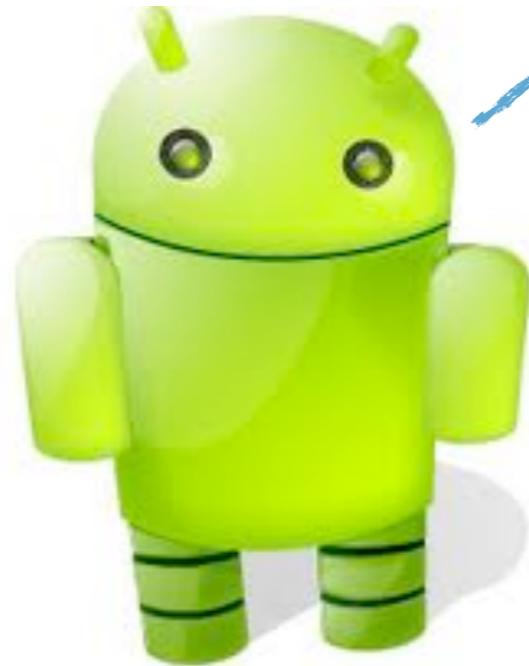
Blog benchmark (Wordpress 3.1)





Thank you !

Questions ?



Talkback

<http://ly.tl/talkback>

Follow-us on Twitter

[@elie](#), [@baptlste](#)