

◦ [October 2006](#) (15)

Stumbleupon mathematics for stumblers

September 19, 2007 — Tim Nash

Give it a [Sphinn](#) | Or a 

What follows is my understanding of the stumbleupon algorithm it is based on some pretty extensive testing using several volunteers however it has been incredibly simplified to make it easier to understand. **We may be totally wrong so just a heads up but I hope this will at least give you some idea of what your thumb up is doing.** I have also written up a few [questions and answers](#) to help people understand what I'm trying to do.

Audience score

Every stumbler has an audience score in the old days stumbleupon told you what your score was but have since taken this facility away. The audience score was based on number of fans, number of pages thumbed up, number of pages thumbed down and number of reviews written. The score is what determines how much stumble juice a person carries.

The audience score has one other factor stumble history. If a stumbler initially stumbles a site and the site receives a large quantity of thumbs up their audience score increases conversely if they initially stumble a site and it's thumbed down their audience score goes down. Stumblers who stumble a site after the initial stumble also have changes to their audience score but not to the same extent.

It is hard to weight which factor is most important when increasing audience score but the factors as I see them are:

- Number of fans
- Number of thumbs up and down you have given
- Stumble thumb bonus - increase to score based on number of thumbs received on a page.

This model means that the obvious technique to get a "power account" is to find more fans, thumb up loads of pages and start stumbles on pages you expect to be popular - sound familiar its pretty much the same on every social media site.

Once we have our idea of an audience score its time to look at a few basic models that stumbleupon might use, you can skip to the big one if you want but these smaller models I think are important to demonstrate individual parts of the algorithm.

A Basic model

Initial stumbler + (number of thumbs up / number of thumbs down) = visitors

This basic model is based on the idea that the initial stumblers audience score will dictate how many visitors will initially see the page and then the number of thumbs up will dictate how many additional people see the page it also presumes that thumbs down have equal weighting to thumbs up.

Audience driven model

Initial stumbler audience + (% of audience of stumbler per thumb up / number of thumbs down) = visitor

This model is a little more complex it presumes that the full audience score is used for the initial stumbler while each additional thumbs up passes a percentage of each stumblers audience score. This model would account for the stumble wave effect, where stumbleupon sends continual waves of varying sizes.

Audience + Domain model

(Initial stumbler audience/#stumbled domain) + ((% of audience of stumbler per thumb up/#stumbled domain) / number of thumbs down) = visitor

This model presumes the number of times the domain is stumbled by a user is a factor therefore the initial stumblers audience score is affected by the number of times they have previously stumbled the domain. If this is done for both the initial stumbler and all stumblers thumbing the page up or down it would explain why mailing lists and friends stumbling the same domain has less and less effect.

The models above show a continual development but there are few more factors rather than showing endless models I will just discuss these factors

Friends

Being friends is not a bad thing while stumbleupon does not provide a bonus it is my belief it does penalise accounts that continue to stumble the same things without being friends or at least one party being a fan. I do not believe the penalty to be huge just a balancing factor to flag that the accounts routinely stumble the same information.

Organic bonus

This I think is a huge factor when a user arrives on a site via the toolbar it is "organic" in the way your arrived, stumbleupon presumes you are judging the page on merits having not seen it before it therefore gives more weight to thumbs up that come via organic stumbling. This is another reason mailing lists fail to work over time on stumbleupon.

Send to

I initially categorised the use of send to as "organic" stumbling but my current belief is that it is not considered organic and therefore does not provide a bonus from organic stumbling, more experiments need to be carried out but I believe it may indeed be the reverse and actually cause a penalty.

The Big one

(Initial stumbler audience /# domain) + ((% stumbler audience /# domain)+ organic bonus - nonfriend) - (((% stumbler audience + organic bonus) + N

So initial stumbler juice is his audience plus his previous stumble bonus which is divided by the number of times the domain has been stumbled by the user. Plus for each thumb up the juice is a percentage of their audience score plus their previous stumble bonus divided by the number of times the domain has been stumbled by that user plus a bonus if the stumble was organic and any to close penalties that may apply. The audience score is reduced by a percentage for each thumb down stumbler plus a bonus if organically stumbled. Finally N which is a random number generator or a Tim get out of jail free card.

The big model is simplified to the extreme but I think is fairly accurate but it does not explain stumble wave suitably so within our model we need to look at time. Sadly we haven't been able to run an experiment beyond a month but based on previous statistical evidence time stumbleupon waves occur on an almost logarithmically with large quantity of waves occurring after the first stumble and then petering out, until the next thumb which sends another series of waves.

Lets follow some examples we will use totally fake numbers to make life easy.

A stumble upon user

Our user lets call him Fred has an audience score of 10 he goes along and starts a new stumble at a site he has never visited it gets a couple of hundred visits and 3 thumbs up

Fred gains a point to his audience score for thumbing something up +a further bonus because others liked his stumble so fred now has an audience score of 13 Fred is really impressed that so many visitors came to his site so he thumbed up another page, even with his increased score it didn't do so well and only 2 people thumbed it up and 2 thumbed it down!

His score is now 14 (increased for thumbing up - no bonus)

Fred tries a different domain it does well and 10 people thumb it up his score goes up to 25, Fred has realised stumbleupon can make him money so thumbs up his proxy site it gets a few visitors but 7 people thumb down the site and 2 marked it as spam. Fred audience score plummets (18 but has been marked by spam so temporarily has his score halved) so his score is now 9 poor Fred will have to work hard to regain his score.

A Domain

Some nice person stumbles the site they had an audience score of 10 which brought a 100 people 3 other people thumbed the site (all came via organic) with scores of 30/100/40 they bring a further 150.

Next day the domain is stumbled again but the number of stumbles is much lower, the owner tries to encourage people to visit the site by using the send to

button and while there are lots of thumbs few extra visitors other than those he sent the send to to.

Mailing lists

A secret group of stumblers have a mailing list, they send an email when they want something stumbled. The first time it worked great and large amount of stumbles followed, the second time it didn't work quite so well soon the mailing list stumbles are counting for little or nothing. (this happens an awful lot repeatedly stumbling the same domain reduces the chance of a stumble wave next time particularly if people outside of the group are not also thumbing up the groups stumbles.

What do you think have we got it right? Wrong? Am I completely off my trolley?

More Stumbleupon content from Tim

Stumbling research

- [Stumbleupon mathematics for stumblers](#) & it's follow up [Questions & Answers](#)
- [What do I & top stumblers have in common](#)
- [Organic Stumbling tip](#)

Stumbleupon vs Digg


- [Page Layouts](#)
- [Voyeurs vs Sheep](#)



[Subscribe to The Venture Skills Blog by Email](#)

All our Posts are audio subscribed for more information [see here](#), and to access the podcast feed [here](#)



 *This blog is moving soon, make sure you move with us by using our Feedburner RSS feed, if you have used the autodiscovery button in your browser you may need to swap feeds, simply delete the old feed and add, <http://feeds.feedburner.com/VentureSkills> For a more detailed explanation on feeds and receiving our content in various formats [click here](#)*

Bookmark 

Posted in [Technology](#), [promotion](#), [social media](#), [stumbleupon](#). Tags: [sem](#). [92 Comments](#) »

92 Responses to “Stumbleupon mathematics for stumblers”

1. [Breaking Down The StumbleUpon Algorithm : The Blog Herald](#) Says: