

Some Thoughts on the TransMail Relevance/Importance Algorithm

A revised version is attached – changes from previous only in section 9.

Because you tend to skim rather than read properly here are the two absolutely central points:

[1] Importance should attach primarily to posts not individuals

Your present system seems to identify TransMail senders simply by name. This works for email but is useless for your system. Whether for incoming or stored email, it is usually the post not the individual that is important. You need to do this either by: (a) Actually communicating with post title in the TransMail or (b) somewhere storing a matrix that associated particular names or TransMail identifiers with posts.

[2] Initial User Inputs Should be More Central

It makes sense for a company or postholder to set *initial* values ie to rank potential known senders to start with. Some could be purely personal local importances (eg wife or friends) but most will relate to the post. There are facilities for changing importance values in your system, but it was not clear to me exactly how easy it would be to set these for initialisation. The user should also be able to set an importance value for someone with whom they initiate a new exchange. Thus eg, when they TransMail to a company asking about wigits, they should be able to indicate on the outgoing TransMail what local importance they wish to attach to the wigit firm. It would be silly for the system to assign a low ranking to the wigit firm because wigit firms are unimportant and because only one TransMail had been sent to them by the user. Presumably, alternatively, a specific reply to a TransMail will automatically have a high local relevance?

You need to read through the words of the section, especially eg 9.1.11 where I have altered it.

Here are some other general comments:

I don't know how central these concepts are to the Bluespace sales pitch, but if important then you really need to allocate more brain-power and resources to it. Again, as I keep saying, you cannot simply assign this to a brilliant mathematician. The mathematics itself may be horrendous, but it is intimately connected with the issues of business use and facility ie what would "make sense" for the system to do.

I also wonder if any consideration has been given to storage space required. The system boasts that this will be reduced because a TransMail is stored only once. But if you have (say) 1000 users, then there are only 1000 global importances but $2 \times 1000 \times 1000$ mutual local importances = 2,000,000. Presumably many of these will be unknown or non existent (unless you store even more information). Is the local importance to have a default value of zero?

A central problem is the misleading nature of the parallel with google. Although both ultimately seek ranking algorithms, there are some crucial differences that mean that starting out with a google model points in wrong directions:

1. When google “gets it wrong” there is no comeback – if the TransMail gives silly answers users will become irritated.
2. Google deals with websites, not individuals, and individual personnel moves are largely irrelevant. TransMail will be radically affected by post much more than individual, yet your system is set up in terms of individuals.
3. Google has no capacity for users to feedback eg to say “Sending me to Kelkoo to buy a dog when all they have is landrovers wasn’t helpful”. Though feedback has now been “bolted on” to TransMail, it should have been central – ie the starting point should have been *what* bits of feedback are wanted *where* rather than adding up numbers of recipients and guessing from what they do with email what they want.
4. For TransMail (and I cannot think of any analogy in google), recipients will get various types of mail:
 - a. Mail from those with whom they are in regular contact [eg PA, line manager, close colleague]
 - b. Mail from those with whom they have initiated new contact [eg they have contacted a company with regard to supplying wigits]
 - c. Mail from friends and relatives [eg spouse – who may or may not work in the same organisation – or friend with some business connections]
 - d. Mail from those in the organisation with whom they do not have regular contact but whose occasional TransMails could be important because they seldom mail [eg facilities management informing of a relevant car park closure]
 - e. Regular (often unread) mail “for information” [eg general circular on vacancies]
 - f. Unsolicited and unwanted mail from outside sources.
5. Global importance will be crucial only if either (a) there is no indication at all of local importance or (b) it has overwhelming upwards effect and the TransMail is one sent to few recipients.

Crucially Google has no means of assigning or asking for “local importance” at present – it cannot rank sites for me when I type in “aardvark” based on my previously analysed interests. It cannot ask me to rank the sites I know of before we start looking – the only reason I am using Google is because I don’t know them. Google therefore has to rely on complex mathematical counts alone. The TransMail system, in contrast, does know about individual preferences and it can specifically ask them. You should therefore *begin* with letting them set the values and only amend these values using your algorithms.

EXHIBIT 2051

metadata slip sheet

BegDoc	REV0000096
BegAttach	REV0000096
File Name	Some Thoughts on the TransMail Relavance.doc
LastCreateOn	5/17/2005 3:39 AM
LastModifiedDate	5/17/2005 8:27 PM
Custodian	Marston, Justin
Author	Marston
SortDate	5/17/2005 8:27 PM
Office-Revision Number	5
Last Saved By	Marston
Exhibit No	