

1  
 ??  
 ??<sup>2</sup>  
 Wikipedia.org1  
 Facebook.comAmazon.com  
 Youtube.comTwitter.com  
 Qq.com  
 Taobao.com

3  
 ??  
 4  
 Big<sup>56</sup>  
 Linked  
 Noisy  
 Unstructured  
 Incomplete

7  
 8  
 9  
 10  
 111213  
 14  
 $(e_i, a_{ij}, s_{ijkl}, h_k, t_l)$   
 $e_i a_{ij} h_k t_l s_{ijkl}$

⋮  
 ??  
 $\tilde{D}S$   
 $\{(d, s) | d \in$   
 $D, s \in$   
 $S\}$

$$(1) \quad g : D \rightarrow S, g(d) = \arg \max_S f(d, s)$$

$gff$   
 15  
 1617  
 ??

$$SD = \sum_{w \in D} S_w * weight(w) * modifier(w) \sum weight(w)$$

$$(2) \quad S_w weight(w) modifier(w) \\ PMI(x, y) = \log_2 F(x, y) F(x) F(y)$$

$$(3) \quad F(x, y) F(x) x$$

$$PMI-IR(x) = \sum_{p \in pWords} PMI(x, p) - \sum_{n \in nWords} PMI(x, n)$$

$$(4) \quad pWords nWords$$

$$\{(f, s_f) | rep(f, ) > \rho_f, s_f = agg(S, f)\}$$

$$(5) \quad f rep(f, ) s_f f agg(S, f) f$$

18  
 ??  
 ??  
 Framework.png

Twitter  
 Twitter  
 Twitter  
 Twitter  
 ??

??  
 1920  
 21

[http://www.cnnic.cn/hlwfzyj/hlwfzxx/qwfb/201408/t20140825\\_47878.htm](http://www.cnnic.cn/hlwfzyj/hlwfzxx/qwfb/201408/t20140825_47878.htm)

[www.alexa.com](http://www.alexa.com)20149

<http://en.wikipedia.org/wiki/Social>

me-  
 dia/

<http://www.twitter.com/>

<http://www.flickr.com/>

[http://www.whitehouse.gov/sites/default/files/docs/big\\_data\\_privacy\\_report\\_may\\_1\\_2014.pdf](http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf)

<http://www.facebook.com/>

[www.amazon.com](http://www.amazon.com)

[www.yelp.com](http://www.yelp.com)

[www.tripadvisor.com](http://www.tripadvisor.com)

[www.twitter.com](http://www.twitter.com)

[www.facebook.com](http://www.facebook.com)

( sentiment