

閲覧数と更新頻度によるスパムブログの傾向分析について

About the trend analysis of the SPAM blog by the number of inspection and the update frequency

田代光輝（たしろ みつてる・Mitsuteru TASHIRO）
ニフティ株式会社／産業技術大学院大学

[Abstract]

Usual blogs and SPAM blogs (Splog) were divided by distributing the spam blog that accounted for 40 percent of a Japanese blog. This paper tries to distinguish SPAM blogs from usual blogs by using the number of inspection and the contribution frequency. Moreover, this result shows the number of inspection of usual blogs becoming lognormal distribution, and obtaining a theoretical average.

[キーワード]

CGM ブログ WEBスパム スパムブログ(splog) ベキ法則 閲覧数 更新頻度

1 はじめに

1.1 ブログについて

Consumer Generated Media（以下：CGM）の発展と普及はインターネット上の発信者・受信者の壁を取り払い、特に情報発信部分におけるコンピューターリテラシーの障壁を低くした。それまである程度コンピューターやネットワーク技術の知識がなければインターネット上での情報発信ができなかった人たちが手軽に情報発信できるようになったことでWEBコンテンツは充実し、インターネットユーザーはより良い情報に多く接する機会を得ることが可能になった。いまやインターネットがメディアの1つ[1]とも言えるようになったのはCGMの普及によるものも大きい。

ウェブログ（以下：ブログ）はCMS(Content Management System)の一種でSNS(Social Networking Service)と並ぶCGMの代表的サービスの1つである。ブログが持つ様々な機能[2]による更新の簡便さや検索エンジンとの相性の良さなどが評価され、誰でも気軽に情報発信できるツールとして普及[3]した。

1.2 WEBスパムとブログ

しかし日本においては価格競争が過度にすすみ無料サービスが普及した[4]こともあり、ブログはWEBスパムの温床ともなっている。2008年3月のニフティの調査[5]によれば日本語のブログのうち約4割がWEBスパムとして利用されているブログ（以下：スパムブログ）である。

WEBスパムの排除はメディアとしてのインターネットの信頼性を上げるためにも重要であり、その抽出手法は様々な形で提案されている。主な手法として”リンク構造による抽出手法[6][7][8]” “言語解析による抽出手法[9]” “更新頻度による抽出手法[10]” の3種類が研究されている（以下：先行研究）。

本研究は先行研究では利用されていない閲覧数を利用し、更新頻度と閲覧数の相関を調査した上で、その結果からスパムブログの効率良い抽出方法の提案を行う。

2 研究のきっかけと仮説

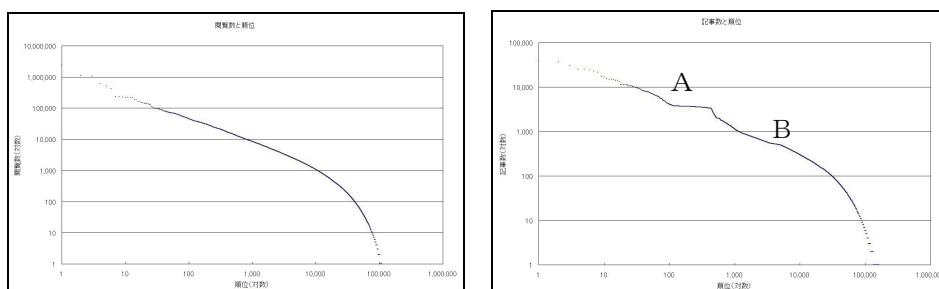
2.1 ベキ法則について

本研究はベキ法則によるニフティの無料ブログサービス「ココログ・フリー」（以下：ココログ）における2008年5月1日～31日までの閲覧数（以下：閲覧数）と2008年5月31日時点の各ブログの記事数（以下：記事数）の解析がきっかけで行われた。このサービスはメールアドレス登録だけで利用できる無料サービスで、2008年5月31日段階で記事数1以上のブログが180,976個ある。日本語ブログ内でのシェア

は 2008 年 3 月の段階で 3.28%[11]である。

ベキ法則とは長く経験則として使われている法則で、ある統計データの値と順位が $y=x^r$ (y =値、順位= $x>0, r<0$) のような分布になることである。x 軸 (順位)、y 軸 (数値) を対数にして散布図にすると直線にちかい分布になるのが特徴である。

図 1 は閲覧数と順位の対数 (10 を底にした対数：以下対数はすべて 10 を底とする) の散布図である。上位から中位にかけてまっすぐ部分と、べき乗グラフに特有の下位の緩やかな「つ」の字のカーブが見てとれる。一方、図 2 は記事数と順位の対数の散布図である。図 1 にくらべ不自然な折れ曲がり (図 2 の A と B) を見ることができる。図 2 中の A・B に該当するブログを実際に確認したところスパムブログであることが確認された。



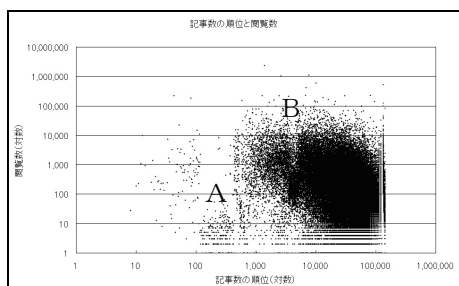
【図 1】閲覧数と順位の散布図 (x 軸=順位の対数 y 軸=閲覧数の対数)

【図 2】記事数と順位の散布図 (x 軸=順位の対数 y 軸=記事数の対数)

2.2 不自然な折れ曲がりの部分について

図 3 は図 2 の順位をそのままに、y の値の記事数から閲覧数に置き換えたものである。上位の部分は記事数が多くても閲覧数が少なく、また図 2 の A と B (図 3 の A と B に同じ) の部分も閲覧数が低くなっているのが見て取れる。下位においても閲覧数が非常に高くなっている部分がみられるなど不自然な分布が確認できる。

この解析からスパムブログと普通のブログは閲覧数を因子として何らかの傾向の違いがあり、それによってスパムブログを効率よく抽出できる手法があるのではないかという仮説を立てた。



【図 3】記事数順位と閲覧数の散布図 (x 軸=記事数の順位の対数 y 軸=閲覧数の対数)

3 ブログ (ココログ) の事前調査

3.1 分類基準

ココログ内の記事数 1 以上のブログ 180,976 個のうち 1,062 個 (母数に対しての抽出条件：誤差範囲 3.00%、想定する調査結果 50.0%、信頼度係数 1.96、以下：条件 A) を抽出し、2009 年 12 月から 2010 年 3 月にかけて目検 (以下：目検調査はすべてこの期間) にて分類した (以下：事前調査)。

ブログの種類はそれぞれ以下のように定義した。

1) 普通のブログ

人間が書いていると判断されるブログ (以下：普通のブログ)。

2) スパムブログ

先行研究においてスパムブログは詳細に分類され様々な分析がされているが、本研究においては機械的に更新されたものやアダルトサイトへの誘導ブログなどと細かく分類せず、一括してスパムブログとして集計した。また既にココログのシステム側でスパムブログとして処理されアクセス禁止となつて内容は確認できないものもスパムブログとして集計をした。

3) 既に削除されているブログ

既に削除されて内容が確認できないブログ（以下：削除済み）。

4) パスワードがかけられているブログ

パスワードがかけられ内容が確認できないブログ（以下：パスワード）。

3. 2 事前調査の結果

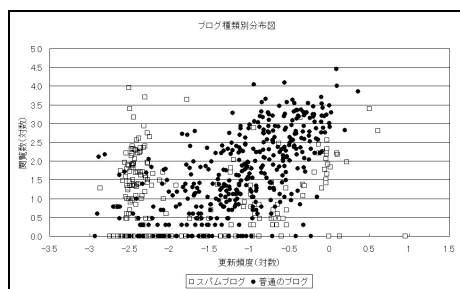
事前調査の内訳は、普通のブログが 552 個で 52.0%、スパムブログが 220 個で 20.7%、削除済みが 251 個で 23.6%、パスワードが 39 個で 3.67%となった。なお、削除済みを母数から除くとスパムブログの割合は 28.5%となる。

〔表 1〕 事前調査の内訳

種類	数	%
普通のブログ	552	52.0%
スパムブログ	220	20.7%
削除済み	251	23.6%
パスワード	39	3.67%
合計	1,062	

普通のブログとスパムブログの「記事数を開設日からから 2008 年 5 月 31 日までの日数で割った更新頻度（以下：更新頻度）」と閲覧数をそれぞれ対数にして散布図にしたのが図 4 である。全体的にレの字型で、 $\log(x)=-2.50$, $\log(y)=0.00$ から $\log(x)=0.00$, $\log(y)=3.00$ あたりにかけて普通のブログが分布し、スパムブログは普通のブログと比較して更新頻度の低い部分で閲覧数が高く、更新頻度が高い部分で閲覧数が低い傾向が明らかとなった。

普通のブログとスパムブログには更新頻度と閲覧数の関係に明確な違いがあることが見てとれる。



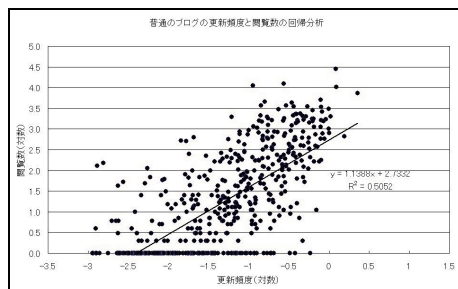
〔図 4〕 事前調査の結果の散布図（x 軸＝更新頻度の対数 y 軸＝閲覧数の対数）

3. 2. 1 普通のブログの傾向

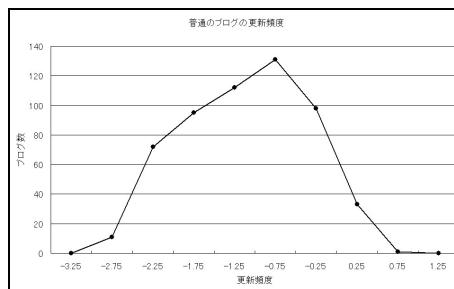
普通のブログは全体的に右上がりの分布になっている（図 5）。

閲覧数は更新頻度に比例して大きくなっており、更新頻度と閲覧数のそれぞれの対数を回帰分析すると、回帰式： $\log(y)=1.14 \times \log(x)+2.73$ 相関係数：0.505 が得られた。

更新頻度を対数で集計すると-1.00～-0.500 の間にピークがくるような分布となる（図 6）。



[図5] 普通のブログの散布図 (x 軸=更新頻度の対数 y 軸=閲覧数の対数) と回帰分析の結果



[図6] 普通のブログの更新頻度の集計 (x 軸=更新頻度 y 軸=ブログの数)

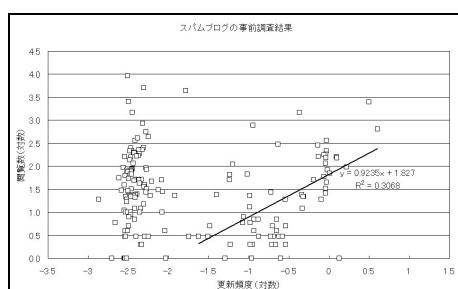
3.2.2 スパムブログの傾向

スパムブログは普通のブログに比べて更新頻度の低い部分の閲覧数が高く、更新頻度が高い部分の閲覧数が低いという傾向がある (図7)。

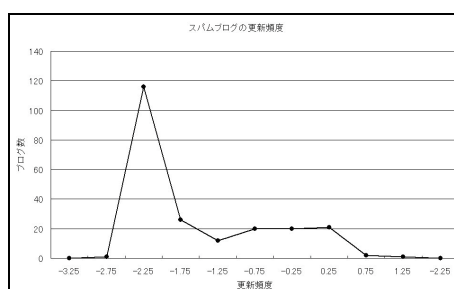
更新頻度の低い部分で閲覧数が高いブログを確認したところ、記事数3以下のブログが92個中91個であったので、閲覧数を記事数3以下と記事数4以上にわけて分析を行った。

閲覧数に関して、記事数3以下ブログでは普通のブログが平均5.04であるのに対してスパムブログが平均213とスパムブログが高い。記事数4以上ブログではやや若干右肩上がりの分布で、閲覧数と更新頻度を対数で回帰分析すると、回帰式： $\log(y) = 0.924\log(x) + 1.83$ 相関係数：0.306 が得られた。

更新頻度を対数で集計すると-2.50前後に大きなピークがある分布となる (図8)。



[図7] スパムブログの散布図 (x 軸=更新頻度の対数 y 軸=閲覧数の対数) と回帰分析の結果



[図8] スパムブログの更新頻度の集計 (x 軸=更新頻度 y 軸=ブログの数)

3.2.4 閲覧数が0のブログの傾向

閲覧数0のものを抽出して分類すると、普通のブログが139個、スパムブログが53個あった。ただし、普通のブログは1個を除いて、残りはすべて4ヶ月以上更新のない放置されたブログであった。

4 追加調査

事前調査ではスパムブログは記事数3以下のブログと記事数4以上で傾向がちがうこと、また記事数4以上であっても閲覧数0のブログは放置されたブログが多い傾向があることから「記事数3以下のブログ」「記事数4以上で閲覧数1以上のブログ」「記事数4以上で閲覧数0のブログ」の3種類に分けて追加調査を行った。尚、各種類のココログ内の数は表2のとおりである。

[表2] 追加調査の対象数

分類	ブログ数	割合
記事数3以下	68,368	37.8%
記事数4以上閲覧数1以上	86,279	47.7%
記事数4以上閲覧数0	26,329	14.6%
合計	180,976	

ココログ内の記事数 3 以下のブログ 68,368 個のうち、1,052 個を（条件 A）目検にて分類した。記事数 3 以下ブログの内訳は、普通のブログが 345 個で 32.8%、スパムブログが 392 個で 37.3%とスパムブログがやや多い。削除済みが 306 個、パスワードが 9 個であった（表 3）。この追加調査で記事数 3 以下のブログは、閲覧数が多くなるにつれてスパムブログの割合が増えることが確認された。

	数	割合
普通のブログ	345	32.8%
スパムブログ	392	37.3%
削除済み	306	29.1%
パスワード	9	0.856%
合計	1,052	

普通のブログは全体の数（345 個）と比較すると閲覧数 10 以上の数が 7.52%（26 個）、閲覧数 32 以上が 3.77%（13 個）と少なくなっていくのに対して、スパムブログは全体の数（392 個）と比較すると閲覧数 10 以上の数が 53.6%（210 個）、閲覧数 32 以上が 39.8%（156 個）となる。

記事数3以下の普通のブログとスパムブログの閲覧数(対数)集計

閲覧数(対数)	スパム	普通
er	121	248
0	41	41
0~0.5	20	20
0.5~1.0	54	13
1.0~1.5	69	8
1.5~2.0	54	5
2.0~2.5	21	1
2.5~3.0	12	0

閲覧数	0～	1～	4～	10～	32～	100～	316～	1,000～
普通のブログ	342	97	46	26	13	5	0	0
スパムブログ	392	271	240	203	156	87	33	12
合計	737	368	286	210	169	91	33	12

この追加調査から普通のブログは更新頻度と閲覧数を対数にすると相関があり、更新頻度は平均-0.931 標準偏差0.517、閲覧数は更新頻度を x としたときに $f(x)=1.15 \times \log(x)+2.94$ を平均として平均-0.931 標

準偏差 0.517 で対数正規分布になっていることが確認された。また、スパムブログは更新頻度・閲覧数に相関がないことも確認された。

さらに、 $f(x)$ 近似した $g(x)=1.15 \times \log(x)+3.00$ と、 $f(x)$ から標準偏差分ずらした $h(x)=1.15 \times \log(x)+2.10$ を基準とすると、「普通のブログの割合が多いエリア」「平均的なエリア」「スパムブログの割合が多いエリア」に分けることが出来ることがわかった。

〔表 5〕記事数 4 以上で閲覧数 1 以上のブログの内訳

	数	割合
普通のブログ	712	67.2%
スパムブログ	154	14.5%
削除済み	145	13.7%
パスワード	48	4.53%
合計	1,059	100%

4.2.1 更新頻度の傾向について

普通のブログの更新頻度の対数は平均-0.931 標準偏差 0.517 で分布している。これを統計解析ソフト SPSS(以下: SPSS)を利用して正規検定を行ったところ、Kolmogorov-Smirnov 検定では有意確率 0.162、Shapiro-Wilks 検定では有意確率 0.006 が得られた。標本数が 2000 以下での検定は Shapiro-Wilks 検定が優位であるために更新頻度は対数正規分布であるとはいえないが、対数正規分布に似た分布になっているといえる(図 10・図 11・図 12)。

スパムブログの更新頻度の対数は平均-0.537 標準偏差 0.590 で分布している。釣鐘状にみえるが、-0.900 と 0.00 の近辺に 2 つのピークのある分布で対数正規分布ではない。



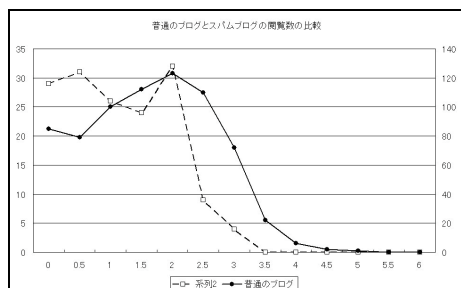
左: [図 10] 普通のブログとスパムブログの更新頻度 (x 軸=更新頻度の対数 y 軸=普通のブログの数)

中: [図 11] 普通のブログの更新頻度の正規検定結果

右: [図 12] 普通のブログの更新頻度の正規 Q-Q プロット

4.2.2 閲覧数の傾向について

普通のブログの閲覧数の対数は、平均 1.87、標準偏差 1.03、スパムブログの閲覧数の対数は、平均 1.39、標準偏差 0.840 であった。尚、閲覧数の 2 の付近にあるピークは図 2 の A のスパムブログの群である。



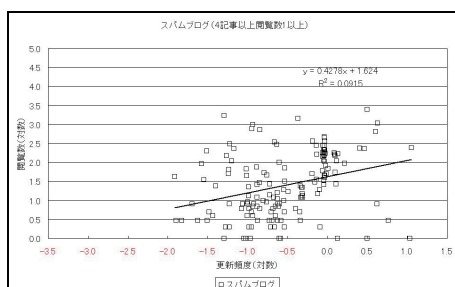
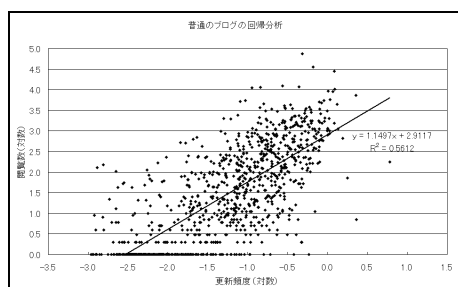
〔図 13〕普通のブログとスパムブログの閲覧数の集計

4.2.3 更新頻度と閲覧数の関係について

普通のブログは更新頻度と閲覧数を対数にすると相関があり、閲覧数は更新頻度に対して対数正規分布をしていることが確認された。また、スパムブログには相関がないことが確認された。

普通のブログの更新頻度を x と閲覧数を y とし、それぞれ対数で回帰分析（記事数 4 以上で閲覧数 1 以上の普通のブログ 712 に加え、記事数 3 以下で閲覧数 1 以上のもの 275 個を別途追加して普通のブログの全体として分析）を行ったところ、回帰式： $1.15 \times \log(x) + 2.91$ 、相関係数：0.561 が得られた。普通のブログでは更新頻度と閲覧数の対数に相関があるといえる。

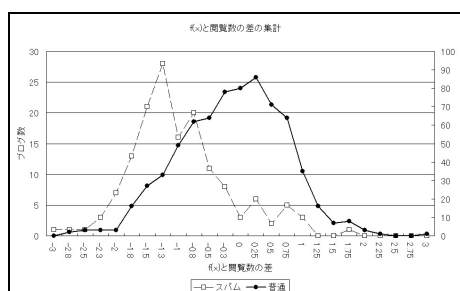
スパムブログの記事数 4 以上で閲覧数 1 以上のブログの回帰分析を行ったところ、回帰式： $0.427 \times \log(x) + 1.62$ で相関係数：0.0915 が得られた。スパムブログでは更新頻度と閲覧数の対数に相関があるとはいえない。



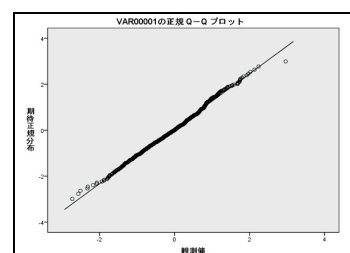
左 [図 14] 普通のブログの散布図 (x 軸=更新頻度の対数 y 軸=閲覧数の対数)
右 [図 15] スパムブログの散布図 (x 軸=更新頻度の対数 y 軸=閲覧数の対数)

さらに記事数 4 以上閲覧数 1 以上のブログに関して、更新頻度を x として $1.15 \times \log(x)$ の値と閲覧数の差を集計し(図 16) 対数にした上で SPSS にて正規性の検定を行ったところ、平均 2.94 標準偏差 0.835、Kolmogorov-Smirnov 検定で有意確率 0.057、Shapiro-Wilks 検定で有意確率 0.147 が得られた(図 17 図 18)。閲覧数は更新頻度に対して $f(x) = 1.15 \times \log(x) + 2.94$ を平均として標準偏差 0.835 の対数正規分布になっているといえる。

スパムブログも同様の集計をしたところ、平均-0.984 標準偏差 0.903 となった(図 16)。



記述統計			
	平均値	標準偏差	標準誤差
N	712		
平均値	-0.0551		
分散	0.6964		
標準偏差	0.835		
最小値	-2.73		
最大値	2.94		
範囲	5.67		
四分位	0.89		
標準誤差	0.032		
信頼区間(95%)			
下限	-0.12		
上限	0.02		



左 [図 16] $f(x)$ の値と閲覧数との差の集計
中 [図 17] 普通のブログの図 16 の分布の正規検定結果
(補足：集計時に $1.15 \times \log(x) + 3$ からの差を集計したため平均が-0.0551 となっている)
右 [図 18] 普通のブログの図 16 の分布の正規検定 Q-Q プロット

4.2.4 分布傾向を利用した分類手法の提案

4.2.3 で得られた回帰式を利用すると以下のような 3 つのエリアに分けることが可能である。

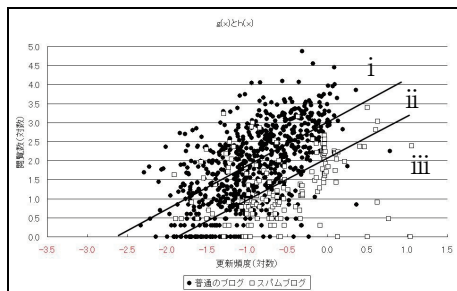
更新頻度を x 、閲覧数を y としたときに $f(x)$ を近似した $g(x) = 1.15 \times \log(x) + 3$ と、 $f(x)$ から標準偏差分だけ離れている $h(x) = 1.15 \times f(x) + 2.10$ を境界線とすると(図 19)、

- i) $\log(y) > g(x)$ には普通のブログの 49.9% (355 個) とスパムブログの 11.7% (18 個) が存在
- ii) $g(x) > \log(y) > h(x)$ には普通のブログの 34.1% (243 個) とスパムブログの 31.2% (48 個) が存在

iii) $h(x) > \log(y)$ には普通のブログの 16.0% (114 個) とスパムブログの 57.1% (88 個) が存在となる(表 6)

ii) に含まれるスパムブログの割合は 16.5% で全体のスパムブログの割合 14.5% (表 5) とほぼ同数であるが、i は 4.83%、iii は 43.6% (表 6) で、全体と比較すると 32.1% と 4.89 倍である。i と iii では 15.2 倍の差となった。

記事数 4 以上で閲覧数 1 以上のブログは $g(x) = 1.15 \times \log(x) + 3.00$ と、 $h(x) = 1.15 \times \log(x) + 2.10$ を基準として「普通のブログの割合が多いエリア」「平均的なエリア」「スパムブログの割合が多いエリア」に分けることが出来るといえる。



[図 19] $g(x)$ 上と $h(x)$ 下のイメージ

[表 6] 記事数 4 以上で閲覧数 1 以上のブログの i・ii・iii の数と割合

	普通のブログ		スパムブログ		スパムブログの割合
	数	割合	数	割合	
$\log(y) > g(x)$	355	49.9%	18	11.7%	4.83%
$g(x) > \log(y) > h(x)$	243	34.1%	48	31.2%	16.5%
$h(x) > \log(y)$	114	16.0%	88	57.1%	43.6%
合計	712		154		17.8%

4.3 記事数 4 以上で閲覧数 0 のブログの追加調査

ココログ内の記事数 4 以上で閲覧数 0 のブログ 26,329 個のうち 1,026 個(条件 A)を目検にて分類した。ココログは仕様として閲覧数 0 が 3 ヶ月続くと削除されるため大半が削除済みである。普通のブログは 511 個と多いが、498 個が 4 ヶ月以上更新されていないブログであり価値ある情報はほとんどないといえる。

[表 7] 記事数 4 以上で閲覧数 0 のブログの内訳

	数	割合
普通のブログ	511	49.8%
スパムブログ	81	7.89%
削除済み	403	39.3%
パスワード	31	3.02%
合計	1,026	

5 まとめ

事前調査および追加調査の結果から、以下のように普通のブログとスパムブログの傾向分析の報告と効率の良い抽出方法の提案を行う。

5.1 更新頻度と閲覧数の相関について

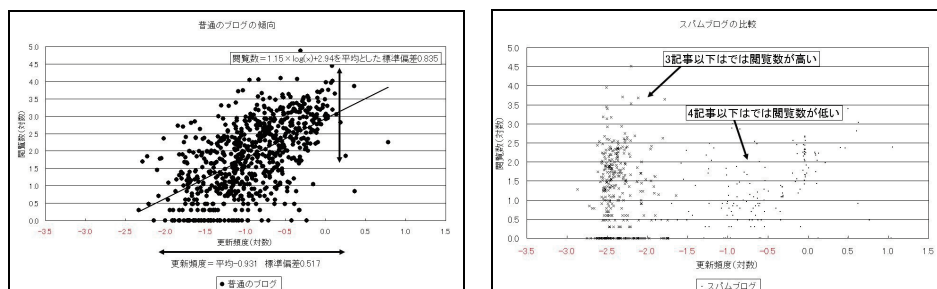
5.1.1 普通のブログについて

普通のブログの閲覧数と更新頻度の対数には相関がある（回帰式： $\log(y)=1.15 \times \log(x)+2.91$ 、相関係数：0.561）。さらに記事数4以上で閲覧数1以上では、更新頻度の対数は平均-0.931 標準偏差 0.517、閲覧数の対数は更新頻度 x に対して $1.15 \times \log(x)+2.94$ を平均とした標準偏差 0.835 の対数正規分布になっている。

5. 1. 2 スпамブログについて

更新頻度と閲覧数に相関はない。

閲覧数は、記事数3以下では普通のブログ（平均 4.53）よりも高い（平均 320）。逆に記事数4以上で閲覧数1以上のものは普通のブログよりも低く、対数にすると更新頻度 x に対して $1.15 \times \log(x)+2.02$ から標準偏差 0.903 で分布している。



左：[図 20] 普通のブログの分布イメージ（x 軸＝更新頻度の対数 y 軸＝閲覧数の対数）
右：[図 21] スпамブログの分布イメージ（x 軸＝更新頻度の対数 y 軸＝閲覧数の対数）

5. 2 効率のよい抽出方法の提案

5. 2. 1 記事数3以下のブログの傾向と分類手法の提案

記事数3以下のブログは、閲覧数が高くなるとスパムブログの割合が多くなる。

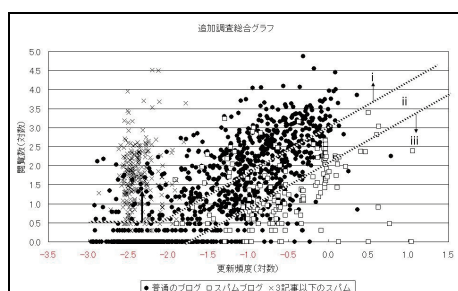
閲覧数32以上になるとスパムブログの占める割合が全体の割合よりも約10倍高くなるため、この条件でブログを抽出すると効率よくスパムブログを抽出できる。（図22）

5. 2. 2 記事数4以上で閲覧数1以上のブログでの分類手法の提案

普通のブログの回帰式と標準偏差を利用すると、以下のような3つのエリアに分けることが可能である。更新頻度 x 、閲覧数 y 、 $g(x)=1.15 \times \log(x)+3.00$ $h(x)=1.15 \times \log(x)+2.10$ とすると、

- i) $\log(y) > g(x)$ は「普通のブログの割合が多いエリア」
- ii) $g(x) > \log(y) > h(x)$ は「平均的なエリア」
- iii) $h(x) > \log(y)$ は「スパムブログの割合が多いエリア」

i と iii を比較するとスパムブログの割合は約15倍変化するため、i では普通のブログが、iii ではスパムブログが効率よく抽出できる（図22）。



[図 22] 抽出方法の散布図（x 軸＝更新頻度の対数 y 軸＝閲覧数の対数）

5. 2. 3 記事数4以上で閲覧数0のブログについて

普通のブログが半数近く存在するがほとんどが4ヶ月以上放置されたブログであり価値のある情報は少ないとみなせる。

本研究の数値・傾向はココログから導き出されたものであるため、他のサービスに応用する際は閲覧数の集計方法等による数値の違い、更新方法やユーザー層の違いによる更新頻度の数値の違いを考慮する必要があるが、分布の傾向は本研究と同様であると推測される。本研究を参考にした他サービスの比較分析が出来ることで幅広い応用の方法が提案されることを期待する。また、先行研究とあわせ本研究において提案する手法によりスパムブログが排除されることでインターネットがメディアとしての重要性を増し、来たる智場社会[12]を支える社会インフラの1つとして貢献することを期待したい。

6 謝辞

本研究にあたり、アドバイスをいただいた東京大学生産技術研究所の豊田准教授、産業技術大学院大学の中鉢准教授、石橋さんをはじめ、分析を手伝っていただいたニフティの松井さんや中野さんなどの社員の皆様に改めて感謝申し上げます。

引用および参考文献

文中の参考文献

- [1]総務省「情報通信白書 平成21年度版」40～41 ページ
- [2]野村総研ブログ調査チーム/シックス・アパート「ブログ白書 2007」3～7 ページ ブログの定義(2006年12月)
- [3]総務省「ブログ・SNS (ソーシャルネットワーキングサイト) の現状分析および将来予測」より、ブログ利用人数は782万人(2005年5月)
- [4]野村総研ブログ調査チーム/シックス・アパート「ブログ白書 2007」124 ページより無料サービス利用率92.3% (2006年12月)
- [5]ニフティ株式会社ニュースリリース「スパムブログのフィルタリング技術を開発」より(2008年3月)
- [6]東京大学生産技術研究所 豊田正史「検索エンジンスパムとその対策技術」(人工知能学会 23 巻 6 号 760～766 ページ 2008年11月)
- [7]神戸大学 佐藤翔平 関和広 上原邦昭 「リンク構造とコンテンツを相補的に用いた極少訓練事例からのスブログ発見」(DEIM Forum 2009)
- [8]島根県立大学 石田和成 「スパムブログの推定と抽出」(日本データベース学会論文誌 Vol.6 NO4 37～40 ページ)
- [9]ニフティ株式会社 高橋哲朗 野田雄也/株式会社富士通研究所 岩倉友哉「スブログの調査と実システムにおける判別手法」(言語処理学会 2008年3月)
- [10]Yu-Ru Lin Hari Sundaram Yun Chi Junichi Tatemura Belle L. Tseng 「Splog detection using self-similarity analysis on blog temporal dynamics」(ACM February 2008)
- [11]ブログファンのブログサービス比較よりアクティブなブログ数の全体とココログの比較より(全体:ココログ=2,273,149:74,563) より (<http://www.blogfan.org/> 2008年3月)
- [12]公文俊平「情報社会学序説 ラストモダンの時代を生きる」(NTT出版 2004年10月)

参考文献

- [1]蓑谷 千鳳彦「統計学のはなし」(東京図書 1997年9月)
- [2]涌井良幸・涌井貞美「図解でわかる 統計解析用語辞典」日本実業出版社 2003年12月
- [3]芳中隆幸 福原知宏 増田英孝 中川裕志「Splog 空間における定量的調査支援システムの開発とその評価」(電子情報通信学会 2008年)

ⁱ e-mail:mitsu_ts@nifty.com