全世界の Web サイトの言語分布と 日本語を含む Web サイトのリンク・地理的位置の解析

童　芳† 　　平手　勇宇†, ††　　山名　早人†, †††
† 早稲田大学大学院基幹理工学研究科 〒169-8555 東京都新宿区大久保 3-4-1
†† 早稲田大学メディアネットワークセンター 〒169-8050 東京都新宿区戸塚 1-104
††† 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-1
† E-mail: {ygmstf, hirate,yamana}@yama.info.waseda.ac.jp

あらまし　Web サーバからは，膨大な情報が発信され続けており，我々の調査では 2006 年 11 月時点で世界中に 537 億のページが存在すると予測されている．これまで，このような大規模な Web を対象としたリンク解析，テキスト解析等が幅広く行われている．しかし，全 Web ページ数のほぼ 13%を占める日本語 Web ページに対しての詳細な解析は存在していない．本稿では，2004 年 1 月～2006 年 8 月の間に収集された 107 億 Web ページに対し， 言語分布，TLD 分布等の解析を行うと共に，2006 年 9 月以降に収集された日本語 Web ページを 1 ページ以上含む Web サーバから発信される約 3 億の Web ページに対して，Web サーバの地理的な位置を特定し地理上での分布・リンク特徴抽出を行った．その結果，日本語で記述された Web ページの 1/3 が jp ドメインであり，他の 2/3 は com 等の別のドメインに存在していることが分かった．日本語 Web ページを発信する Web サーバについては，その 3/4 が日本国内に存在するが，他は国外に存在していることが分かった．

キーワード　Web マイニング，リンク解析，TLD と言語，地理的な位置

Analysis in Language Distribution of World Web Sites and Links/Geographic Positions of Web Sites including Japanese

Hou TOU† 　Yu HIRATE†, ††　and　Hayato YAMANA†, †††
† Graduate School of Fundamental Engineering, Waseda University 3-4-1 Okubo, Shinjuku-ku, Tokyo, 169-8555 Japan
†† Media Network Center, Waseda University 3-4-1 Okubo, Shinjuku-ku, Tokyo, 169-8555 Japan
††† National Institute of Informatics 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430 Japan
† E-mail: {ygmstf, hirate,yamana}@yama.info.waseda.ac.jp

Abstract　According to the investigation result in Nov 2006, the number of web pages all over the world is estimated 53.7 billion. In order to extract useful information from such a great number of web pages, various web analysis methods, such as link analysis and text analysis, are being widely applied. However, in previous statistical analysis of web, most are based on the whole world web sites. But a more detailed analysis of Japanese web pages which account for almost 13% of the number of the whole world web pages, has never been done before. In this paper, we made a statistical analysis of web pages including Japanese crawled after Sep 2006.

Keyword　web mining, link analysis, TLD-Language, geographic position

1. はじめに

近年，Webサーバから発信される情報量が膨大になり，2006 年 11 月の時点で世界中に 537 億のページが存在するという調査結果が報告されている[1]．このような大量のWebページから有用な知識を抽出するために，リンク解析をはじめとするWebマイニングに関する研究[2][3][4][5][6][7]が幅広く行われている．Webマイニング手法を適用するにあたり，ドメインごとのページ数，言語分布，Webページ間のリンク傾向等，Webに関する統計情報を大規模に把握することができれば，さらに有益なマイニング結果を得ることができると考えられる．

従来の研究・調査では，世界中のWebページを対象に行うものが多い．2000 年 7 月に，International Software Consortium（ISC）がインターネットに接続しているホストのドメイン名の分布状況について統計調査を行っている[8]．その調査結果によると，第 1 位のcomと第 2 位のnetだけで、全体の 60%強を占め，jpドメイン数は 4 位で全体の 3.67%にとどまる．また，2006 年に加藤真らはe-Societyプロジェクト[10]によって収集された世界中 120 億ページの内の 30 億ページに対して統計調査を行っている[11]．その統計結果によると，Webサーバのドメインの分布状況はcomドメインの圧倒的な数が見られている一方, jpドメイン数が 2 位になることを示している．

また，Webページのドメイン分布状況を見てみると, 2001 年に池内ら[9]の調査結果より,jpドメインはcom, net, eduに続いて 4 位であり, 全体 5%未満である．一方，2006 年加藤真らの調査結果[11]により，その分布状況はcom, org, netに続いて，netと若干の違いで 4 位となることがわかった．

一方，Web ページの言語の分布については，2001 年 2 月池内らと 2006 年加藤真らの調査結果により，5 年間内で日本語 Web ページ数が全体の 5%程度から全体の 13%に量が増え，ドイツ語 Web ページ数を超え 2 位となることが示されている．

以上示したように，ここ 6 年間で日本語 Web ページ数の量が著しく増え，世界中の Web ページ全体の構成に対してますます重要な一部分になることがわかる．本研究では，2004 年 1 月から 2006 年 8 月までに収集された世界中の 107 億 Web ページに対し，分布統計手法を適用して，最新の統計結果を示す．

また，本稿では，日本語を 1 ページ以上含む Web サーバから 2006 年 9 月に収集された約 3 億の Web ページを対象に，Web サーバの地理的位置の分布を明らかにする．

本稿では，次のような構成をとる．まず，第 2 節では既存の統計情報と地理位置の関連研究について述べる．次に，第 3 節でデータセットと具体的な解析方法について述べる．第 4 節では各統計結果を示した後，第 5 節でまとめを行う．

2. 関連研究

2.1 International Software Consortium（ISC）による調査

ISCは 2007 年 7 月に，インターネットに接続している 21 億ページのホストのドメイン名の分布状況について調査を行っている[8]．この調査結果では，1 位のcomに属するホストは全体の 35.14%を占め，32,696,253 ホストであった．2 位は全体の 25.18%を占めるnetに属するホストで, ホスト数は 23,432,135 であった．この 2 つのドメインだけで全体の 60%のホストを占める．続いて 3 位はeduの 6,678,055 ホストで 7.18%を占め，4 位はjpドメインの 3,413,281 ホストでわずか全体の 3.67%と報告されている．また，国別ドメインと言語分布の統計調査も実施されており，英語圏の国が圧倒的なドメイン数で上位にランクされている．一方， jpドメイン数及び日本語は，ともに 4 位になっている．

2.2 池内らによる調査

2001 年 2 月に池内らはISCの調査結果に対してWebページ数と言語分布について統計調査を行った[9]．その結果によると，Webページ数の分布は，ISCが統計したホストのドメイン分布結果とほぼ一致したが，一部のドメインに分布の偏りが存在したと報告している．ホスト数に比べてWebページ数の比率が多いドメインとしてcom, org, de, gov, ruなどが挙げられた．一方，ホスト数が多いにも関わらず，Webページ数が少ないドメインとしてnet，milなどがあった．jpドメインの場合は，ホスト数とWebページ数のいずれの結果でも，全体の 3~4%を占めると報告されている．また，言語分布結果によると，自国ドメインのページの比率よりも自国語のページの比率の方が高い現象が指摘されている．

2.3 加藤真らによる調査結果

加藤らは，2004 年 1 月から 2006 年 7 月末までに収集したWebページ中の約 30 億のWebページに対してWeb構造を中心に解析を行った[11]．この解析結果によると，WebサーバとWebページのドメイン分布において，comが圧倒的な量で 1 位に占める結果となった．一方，Webサーバ数の 2～4 位は，順にjp，org，itとなり，ページ数の 2～4 位は順にorg，net，jpの順となり，いずれも[8]とは違った結果となった．すなわち，インターネットに接続するホスト数とWebサーバ数の分布は比例しない．

また，言語分布の分析では，日本語 Web ページ数が全体の 13%を占め，2 位であることを示している．他の統計項目には，言語－トップレベルドメイン（TLD）分布，TLD－TLD 分布などがあり，これらの解析結果により，世界中の Web ページの中では日本語 Web ページの強い強連結なページ群の巨大化傾向を示している．

2.4 Web ページの IP アドレスで地理位置を特定

近藤らの研究[15]では，URLをIPアドレスに変換してWebページが属するホストの地理的な位置を特定している．[15]では，任意のページにリンクしているWebページ集合の地理的位置を特定し，地理的位置の分散状況によって当該ページの地域性を評価している．近藤らは，IPアドレスをランダムに 100 万件を生成して調べた結果，アメリカが 1 位で約 58.17%，日本が約 2 位で 6.06%である結果を報告している．

2.5 関連研究のまとめ

2.1～2.3 の調査結果により，2000 年から 2006 年までの 5 年間内で、日本語 Web ページの全 Web ページに占める割合が大きく変化していることが分かる．また，2.4 では地域の特定により，Web ページ情報を直観的に把握でき，Web ページの地域性を評価できることが示されている．

本研究では，以上の従来の研究を踏まえ，規模をより大規模にし，世界中の 107 億ページに対して解析を行った．これによって，最新の Web ページの統計結果を示す一方，日本語 Web ページの Web 上及び地理上の分布・リンク特徴を抽出する．また，日本語 Web ページを中心とする 3 億のページに対して地理位置分布・リンク情況の解析を行うことにより，より深く特徴の抽出を目指す．

3. 解析方法

3.1 対象とする Web ページデータセット

解析対象のWebページは，e-Societyプロジェクト[10]によって収集されたWebページを対象とする．e-Societyプロジェクトでは，2004 年 1 月からWebページの収集を継続している．2004 年 1 月～2006 年 7 月は，全世界のWebページの収集を行い合計で 144.5 億のWebページを収集完了している．一方，2006 年 9 月からは，144.5 億のWebページをもとに日本語ページを 1 ページ以上持つWebサーバ約 140 万サーバを対象に毎月更新収集を行っている．

本稿では，全世界のWebページの解析には，2006 年 7 月までに収集された 144.5 億Webページの中から 10,696,996,553 ページ[1]を対象とし，日本語のページの解析には，2006 年 9 月に収集された 303,174,643 ページを対象とした．

3.2 解析方法

3.2.1 解析対象とする Web ページ情報

Web の傾向調査のために，全ての Web ページに対し，次の 3 つの情報を抽出し，解析の対象とした．

1. TLD 情報

IANA[16]が 2007 年に公開した標準となる 272 個 TLD名によって，収集されたページが属するページを判定した．なお，ドメイン名が割り当てられていないIPアドレスについては，「その他」として分類を行った．

2. 言語情報

我々は，Webページの収集時に当該ページの言語判定を行っている．言語判定には，ベイシスの言語判定システム[17]を利用している．判定できる言語は，英語，日本語，中国語，フランス語，韓国語，スペイン語，ドイツ語，イタリア語，ロシア語，ポルトガル語，アラビア語の 12 言語である．また，画像，動画等は，バイナリとして判別される．12 言語に判別できなかった言語は，その他「oth」として分類される．

[1]実際には約 144.5 億ページを収集したが Disk 故障等によりデータとして利用できる 107 億ページを利用した．

3.　経度・緯度情報

任意の Web ページの地理的位置を特定するために，まず Web ページから Web ページが配信されている Web サーバの IP アドレスを特定する．次に，IP2Location™社の IP－経度・緯度－国名・市名変換テーブル[18]を用いて，IP アドレスから経度・緯度情報への変換を行った．

3.2.2　Web 解析項目

我々は，Web の統計情報を計算するために，次に示す 4 つの項目について解析を行った．

1.　TLD 毎の言語分布

任意の TLD に含まれる Web ページが，どの言語で書かれているのかを把握するために，すべての TLD の言語分布の調査を行った． この解析は，全世界の約 107 億ページに対して実施を行った．

2.　リンク元とリンク先の TLD 分布

TLD 間でのリンク関係を調査するために，データセットに含まれているすべてのリンクを，リンク元の TLD とリンク先の TLD のペアによって分類を行った．この解析は，日本語を中心とする約 3 億ページに対して行った．

3.　地理的位置の TLD 分布

全世界を，経度 5 度，緯度 5 度四方セルに分割を行い, 任意のセルから配信されている Web ページが，どの TLD に属するのかの調査を行う．この解析は，2006 年 9 月に収集された日本語ページを 1 ページ以上含む Web サーバから発信される約 3 億ページに対して行った．

4.　地理的位置の言語分布

前述の項目と同様に，任意のセルから配信されている Web ページがどの言語で書かれているかの調査を行う．本解析は，2006 年 9 月以降毎月収集されている日本語を 1 ページ以上含む Web サーバから発信されるのべ約 19 億ページに対して行った．

4.　解析結果

4.1　全世界の約 107 億ページに対する解析結果

4.1.1　TLD(Top Level Domain)分布

解析対象となる 107 億ページに対し，TLDごとのWebページ数，TLDごとのWebサーバ数（ホスト数）のカウントを行った．TLDごとのWebサーバ数分布を図 1に示す．また，TLDごとのWebページ数分布を図 2に示す．図 1，図 2では，Webサーバ数が多い上位 20 位のTLDと，21 位以下のTLDを”other”としてまとめあげた際の分布を示している．

さらに，図 3はWebサーバ数とWebページ数の比率状況を示している．図 3は，左軸がWebサーバ数（ホスト数）を示し，右軸がページ数を示す．ホスト数の軸とページ数の軸が同一となっている部分では， 1 ホスト 250 ページであることを示す．したがって，折れ線グラフが棒グラフより高い場合は， 1 サーバあたりのページ数が 250 ページを超えていることを示す. また, 図 3の中の各TLDの内訳は表 1の通りである．

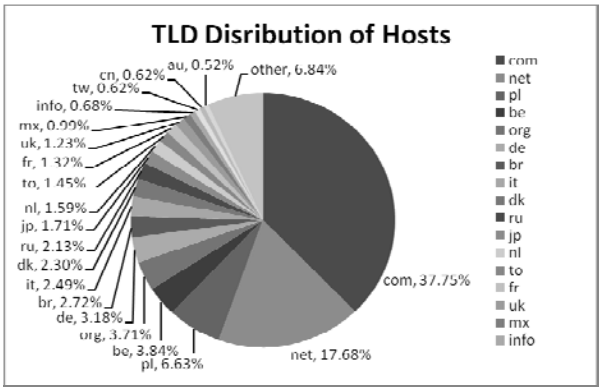

図 1　107 億ページに対応する Web サーバの TLD 分布

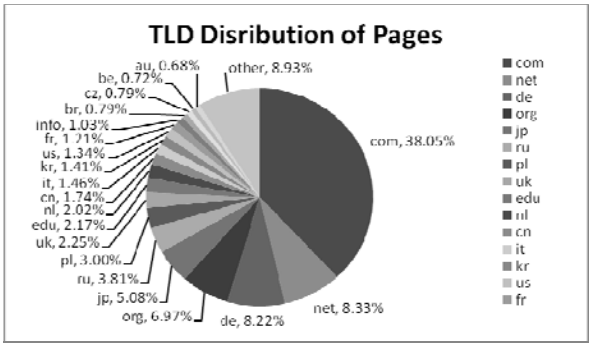

図 2　107 億ページの TLD 分布

全体的に見れば，ドメインによってページの偏りがあることが確認できる．ホスト数に比べてページ数が少ないのは net, pl, be などがある．一方，ページ数が多いのは de, org, edu などが挙げられる．com は Web サーバ数両方とも約 38%で，1 サーバあたりの Web ページ数は平均的な値となっている．一方，jp ドメインの場合は，Web サーバ数の分布率の 1.71%に比べ，ページ分布率が 5.08%であって，ページ数の方が大きい値となった．

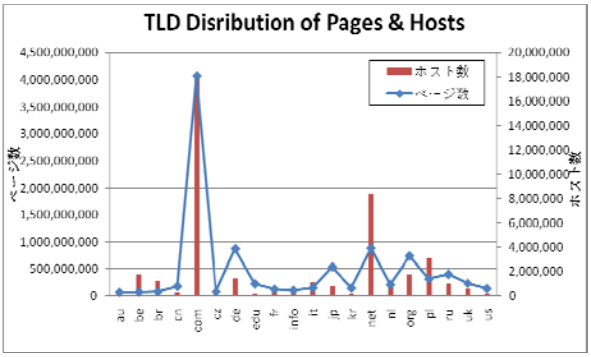

図 3　ホスト・ページの TLD 分布比較

表 1　TLD名対照表（[16]より抜粋）

TLD 名	内訳
au	Australia
be	Belgium
br	Brazil
cn	China
com	Commerce
cz	Czech Republic
de	Germany
edu	postsecondary institutions accredited by an agency on the U.S. Department of Education's list of Nationally Recognized Accrediting Agencies
fr	France
info	Information
it	Italy
jp	Japan
kr	Korea, Republic of
net	Network
nl	Netherlands
org	Organization
pl	Poland
ru	Russian Federation
uk	United Kingdom
us	United States

4.1.2　全世界の言語分布

107 億ページの言語の分布状況を図 4に示す．英語ページは 2006 年 2 月の全体の 2/3 に占める量から 2/5 に減量したことがわかった[2]．また，割合が増加した言語は，ドイツ語および判定不能な「oth」であった．日本語ページは両方とも 13%くらいであり，2006 年 2 月から比率的には変わっていないとは言える．

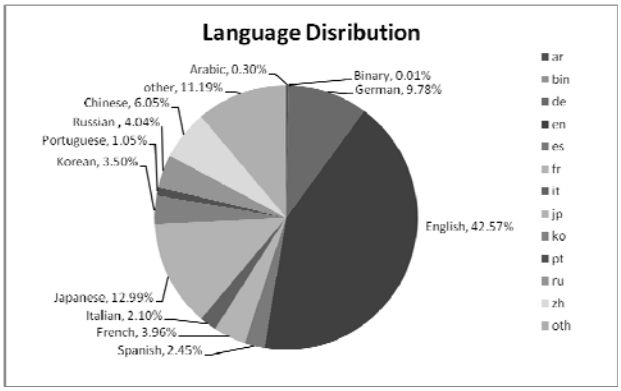

図 4　107 億ページにおける言語の分布

一方，図 5は，日本語を 1 ページ以上含むWebサーバから発信される約 3 億のWebページの言語分布を示す。日本語を 1 ページ以上含むWebサーバの言語分布であるため，日本語が量的には多く，全体の約 50%を占めることを示している．ほかには，日本語ページとリンク関係が強い英語のページが約 32%を占める．その他の言語は，中国語約 7%，判定できない言語約 6%，フランス語約 2%があるなどを示しており，これらの言語分布から日本と他国との関係を読み取ることができると考えられる．

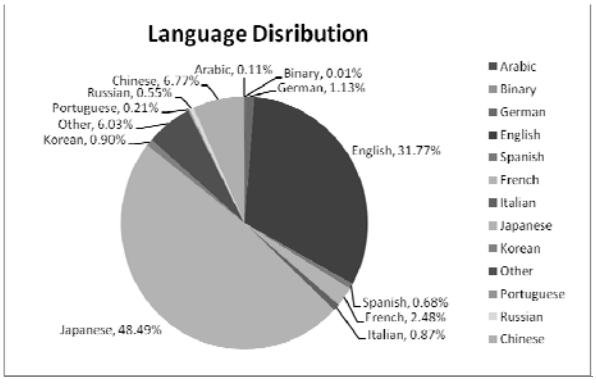

図 5　3 億ページにおける言語の分布

[2] 我々の Web ページの収集は，jp ドメインを起点として収集を行った．したがって，2006 年 2 月時点では，日本語，英語以外のページに十分に到達できていなかったことが予測される．

4.1.3　jp ドメイン内の言語分布

図 6は言語とTLDの関係を示す図である．jpドメインのページ内ではほぼ 90%が日本語ページとなった．英語ページが約 5%である一方，そのほかの言語がどれでも 1%未満となった．つまり，日本国内から発信されたWebページが 9 割以上日本語に書かれたページであるという結論を得た．

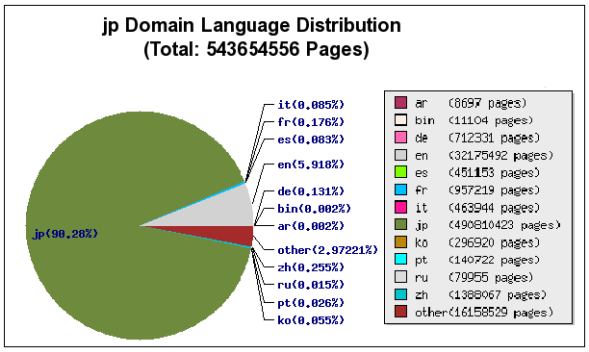

図 6　107 億ページにおける jp ドメインの言語分布

4.1.4　日本語 Web ページの TLD 分布

次に日本語で記述されたWebページのTLD構成について解析した．対象データは 107 億のWebページであり，この中から日本語で記述されていると判定された約 13%のページのTLD分布を図 7 に示す．図 7に示すように，jpドメインよりもインターナショナル・ドメインであるcomのほうが全体に占め量が多い．jpドメインは約 35%で 2 位となった．結論として，日本語Webページの半分以上は地域の明示されていないインターナショナル・ドメインに属して，約 1/3 がjpドメインに存在する．

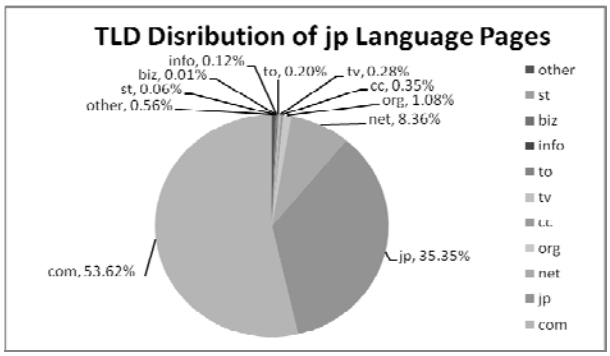

図 7　107 億ページにおける日本語ページの TLD 分布

4.2　日本語を 1 ページ以上含む Web サイトから収集した約 3 億ページに対する解析

4.2.1　TLD ごとのリンク先 TLD の分布

日本語 Web ページを持つ Web サーバから発信される情報に対してより詳しい解析を行うため，日本語 Web ページを 1 ページ以上含む Web サイトから発信される Web ページ間のリンク関係を TLD 単位で調査した．表 2 に，出現頻度の多い順に 10TLD を選択し，アウトリンク先の TLD 分布を示す．

表 2に示す通り，アウトリンク数の 1~3 位はcom, jp, netであるが，どのドメインでも自身と同じドメインへのリンク数が最も多い．しかし，comへのリンク数が比較的多いことも分かる． jpドメインのリンク先のTLD分布では，jpドメインの言語分布の場合とほぼ同じく，jpが 90%に占めるという結果を得た．

表 2　日本語を 1 ページ以上含む Web サイトから収集された約 3 億ページにおける主な TLD ごとのリンク先 TLD の分布

From	アウトリンク数	biz	cc	cn	com	de	info	jp	net	org	pl	other
biz	115,300,153	**73.85%**	0.05%	0.01%	11.60%	0.05%	0.77%	9.19%	2.04%	0.77%	0.10%	1.57%
cc	261,201,539	0.06%	**80.15%**	0.01%	4.48%	0.02%	0.17%	13.05%	1.34%	0.29%	0.04%	0.40%
cn	240,647,284	0.00%	0.04%	**95.13%**	3.57%	0.01%	0.01%	0.02%	0.79%	0.13%	0.00%	0.30%
com	6,841,488,480	0.03%	0.02%	0.15%	**95.39%**	0.09%	0.19%	1.11%	1.16%	0.73%	0.07%	1.07%
de	159,815,863	0.02%	0.01%	0.01%	3.48%	**92.26%**	0.15%	0.04%	0.65%	1.67%	0.04%	1.66%
info	115,410,050	0.10%	0.01%	0.01%	8.29%	0.11%	**86.50%**	0.54%	0.70%	1.72%	0.05%	1.97%
jp	3,942,314,563	0.18%	0.05%	0.11%	4.63%	0.08%	0.51%	**90.61%**	1.81%	0.70%	0.13%	1.18%
net	1,508,257,267	0.07%	0.04%	0.31%	6.66%	0.26%	0.36%	3.01%	**86.80%**	1.02%	0.08%	1.38%
org	489,799,667	0.05%	0.03%	0.12%	5.83%	0.20%	0.79%	1.87%	1.27%	**88.18%**	0.08%	1.60%
pl	492,546,373	0.05%	0.00%	0.01%	2.15%	0.05%	0.32%	0.02%	0.32%	0.87%	**95.43%**	0.79%

4.2.2　Web サーバの地理的な位置分布

日本語Webページを発信するホストのIPアドレス177,054 個(約URLが示すホストの 20%)を特定し，それらをIP－経度・緯度－国名・市名変換テーブル[18]によって地理位置に変換した. その結果を次の図 8に示す.

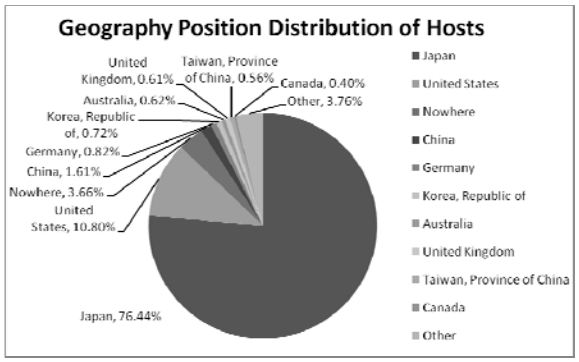

図 8　日本語 Web ページを発信する Web サーバの地理位置分布

図からわかるように，日本語Webページを発信するWebサーバの 3/4 が日本国内に，1/10 がアメリカに分布している．この２つの地域に分布しているホストを合わせて全体の 90%弱を示している．また，日本以外のアジアに分布するWebサーバを含めると全体の 80%以上となる．他には，明確に示されていない地域（nowhere）に分布するホスト数が全体の3.66%を占める．これらの結果は4.1.4でドメインにより示された分布と違っており，TLDと実際にWebサーバが設置される国が大きく異なっていることがわかる.

4.2.3　言語種類の地理的な分布

日本語を 1 ページ以上含むWebサーバから収集した 3 億Webページの言語種類の地理上分布について調査を行った．その結果を図 9に示す.

図からわかるように，発信されたページ数が多く，かつ言語種類が密集している地域が順に日本と中国，アメリカ，ヨーロッパの 3 箇所であることがわかった. 各地域に3.2.1に示す 13 種類の言語が全て揃っている．また，他の地域からは，例えば，アフリカ中央部，オーストラリア，ブラジルなどから発信されたページ数が比較的に少なく，ページの言語種類も共に少なく，平均的に 6 種類になっている.

これらは，元来は日本語ページを中心とする情報発信を行う Web サーバではなく，多言語での発信を主とするが日本語についても翻訳を掲載している等，日本との密接なつながりがある Web サーバであると考えられる.

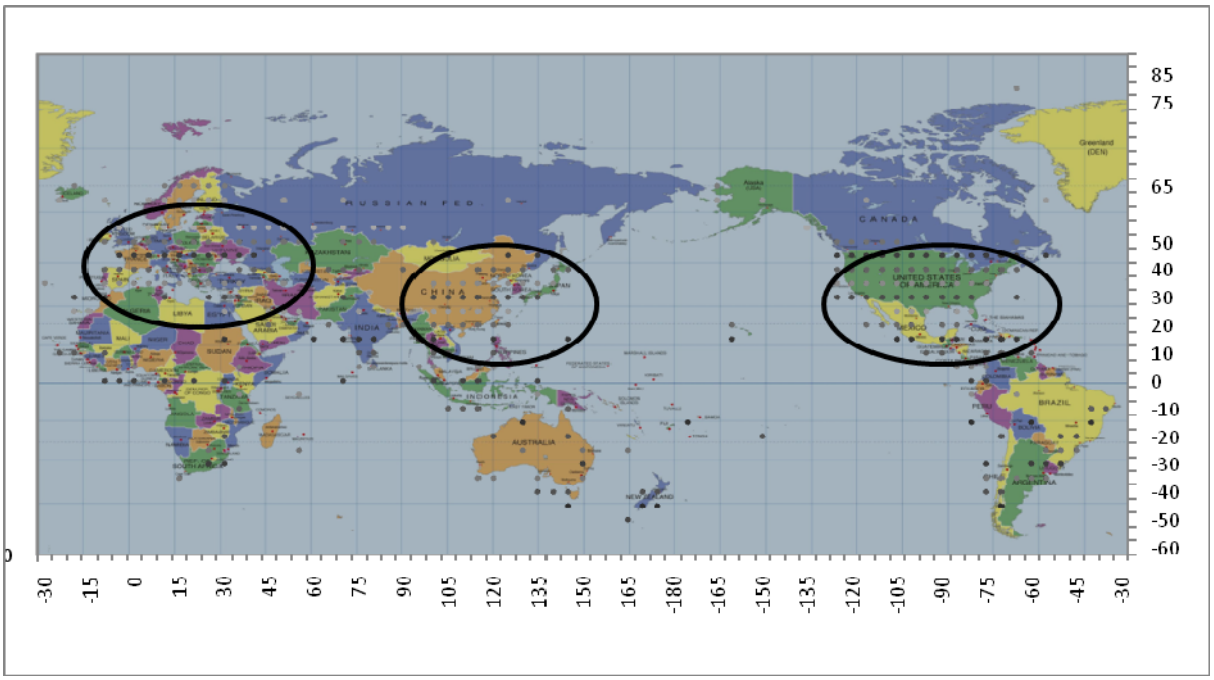

図 9　日本語を 1 ページ以上含むWebサーバから発信されるWebページの言語種類の地理上での分布（地図：デザインエクスチェンジ社のMAPIOシリーズ製品（著作権フリー）の地図[19]を利用）

5. おわりに

本稿では，2004 年 1 月から 2006 年 7 月までに収集された世界中の 107 億 Web ページと，2006 年 9 月に収集された日本語 Web ページを 1 ページ以上含む Web サーバから発信される 3 億の Web ページに対して解析を行った. 統計調査にあたっては, TLD, 言語，地理的な位置の各項目で分布とリンク情報を統計した．その結果，近年，日本語ページ数が増加していることがわかった．また，日本語 Web ページを含む Web サーバの TLD 分布では，約 1/3 が jp ドメインに存在することがわかった．これはすなわち，jp ドメインのみをクローラの収集対象としただけでは，日本語で記述された Web ページの 1/3 しか収集できないことを示す．また，日本語 Web ページを発信する Web サーバの地理的位置を解析したところ約 3/4 が日本国内に存在するとわかった.

謝辞

本研究の一部は、文部科学省リーディングプロジェクト「e-Society」及び情報爆発プロジェクトとして実施した.

参考文献

[1] Y. Hirate, S. Kato and H. Yamana, “Web Structure in 2005”. In Proc. of the WAW2006, 2006.

[2] S. Lawrence, and C. L. Giles, “Searching the World Wide Web”, Science, Vol.280, No.5360, pp.98-100, 1998.

[3] S. Lawrence, and C. L. Giles, “Accessibility of Information on the Web”, Nature, Bol.400, pp.107-109, 1999.

[4] Sepandar D. Kamvar, Taher H. Haveliwala, C. Manning, and G. Golub, “Exploiting the block structure of the web for computing PageRank”, Technical Report, Stanford University, 2003.

[5] G.Flake, S.Lawrence and C. Giles, “Efficient identification of Web communities”, In Proc. of 6th ACM SIGKDD Conf., 2000.

[6] P.K. Reddy and M. Kitsuregawa, “Anapproach to relate the Web communities through bipartite graphs”, In Proc. of 2nd Int. Conf. on Web Information Systems Engineering, 2001.

[7] 村田剛志，“参照の共起性に基づく Web コミュニティの発見”, 人工知能学会論文誌, Vol.16, No.3, 2001.

[8] Internet Software Consortium , http://www.isc.org/

[9] IKEUCHI Home Page, http://www.daito.ac.jp/~ikeuchi/webmetrics/webmetrics_1.html

[10] e-Society プロジェクト，http://www.yama.info.waseda.ac.jp/~yamana/es/

[11] 加藤真，山名早人，Fact of the Web:30 億ページのウェブの解析，DEWS2006.

[12] 平手勇宇，山名早人，全世界の Web ページの TLD・言語分布解析，IPSJ 全国大会 2008.（採録）

[13] 張建偉，石川佳治，北川博之，“空間情報ハブ抽出のためのウェブリンク解析手法の開発”, DBSJ Letters, 2004.

[14] 井上陽介，李龍，高倉弘喜，上林弥彦，“地域情報検索のためのリンク構造分析によるウェブページと地域の関係抽出”, 電子情報通信学会データ工学ワークショップ，2002.

[15] 近藤浩之，手塚太郎，田中克己，“リンク元ページのアドレス情報に基づく Web ページの地域的支持度の分析”，DEWS2007.

[16] IANA: Root-Zone Whois Index by TLD Code, http://www.iana.org/root-whois/index.html

[17] Basis Technology Rosette 言語判別システム，http://www.basistech.co.jp/language-identification/

[18] IP2Location™IP-Country-Region-City-Latitude-Longitude-ISP-Domain Database[DB8]，http://www.ip2location.com/ip-country-region-city-latitude-longitude-isp-domain.aspx

[19] デザインエクスチェンジ社の MAPIO シリーズ製品，http://www.dex.ne.jp/product/deximage/mapio.html