

序章

検索エンジン(Google/Yahoo)が自らのアルゴリズムを公開することはありえないなか、SEO対策(検索エンジン最適化)には、2つのアプローチが存在する。

1つ目は、リバース・エンジニアリング的なアプローチであろう。

検索エンジン・アルゴリズムの設計図を手にするできない中で、起きている事象を分析する中から一定の解を導く、この手法が日本国内のSEOの主流といえる。

そして、2つ目のアプローチとして Google や Yahoo 等の検索エンジン各社が自社の独自技術をライバルに模倣されないために取得する特許や論文を読み解くで、アルゴリズムにせまるアプローチが存在する。英語圏では、このアプローチを前提にリバースエンジニアリング・アプローチで、より客観的分析を行う手法が一般的といえる。

そうした中で、例えば、Google 創業者2人とスタンフォード大学の連名で取得された PageRank に関する特許や HITS 理論や現在の Yahoo アルゴリズムの根幹に当たる TrustRank(SEEDアルゴリズム)に関する論文や特許などは、英語圏では、SEO 対策のあり方そのものに多大な影響を与えてきた。

そして、本レポートで解明していく Google 特許(パテントNo.6,725,259B1)も、英語圏では重要視され、SEO対策のあり方そのものを変えた重要な特許の1つと言える。

なぜなら、より検索アルゴリズムの根源的な領域に属す内容であるリンクスコア(=リンクから獲得する特典・リンク価値)に大きな影響を与える特許であるからである。

少し前置きが長くなったかもしれない。

早速、特許の内容にアプローチしていこう。

自サイトとリンク元のIPアドレスにより算出されるリンクスコア

まずは、リンク元と、そのリンクを受けているサイト(自サイト)、それぞれのIPアドレスによりどのようにリンクスコアが算出されるのかについて。

特許(パテントNo.6,725,259B1)から抜粋した次の図1をご覧ください。

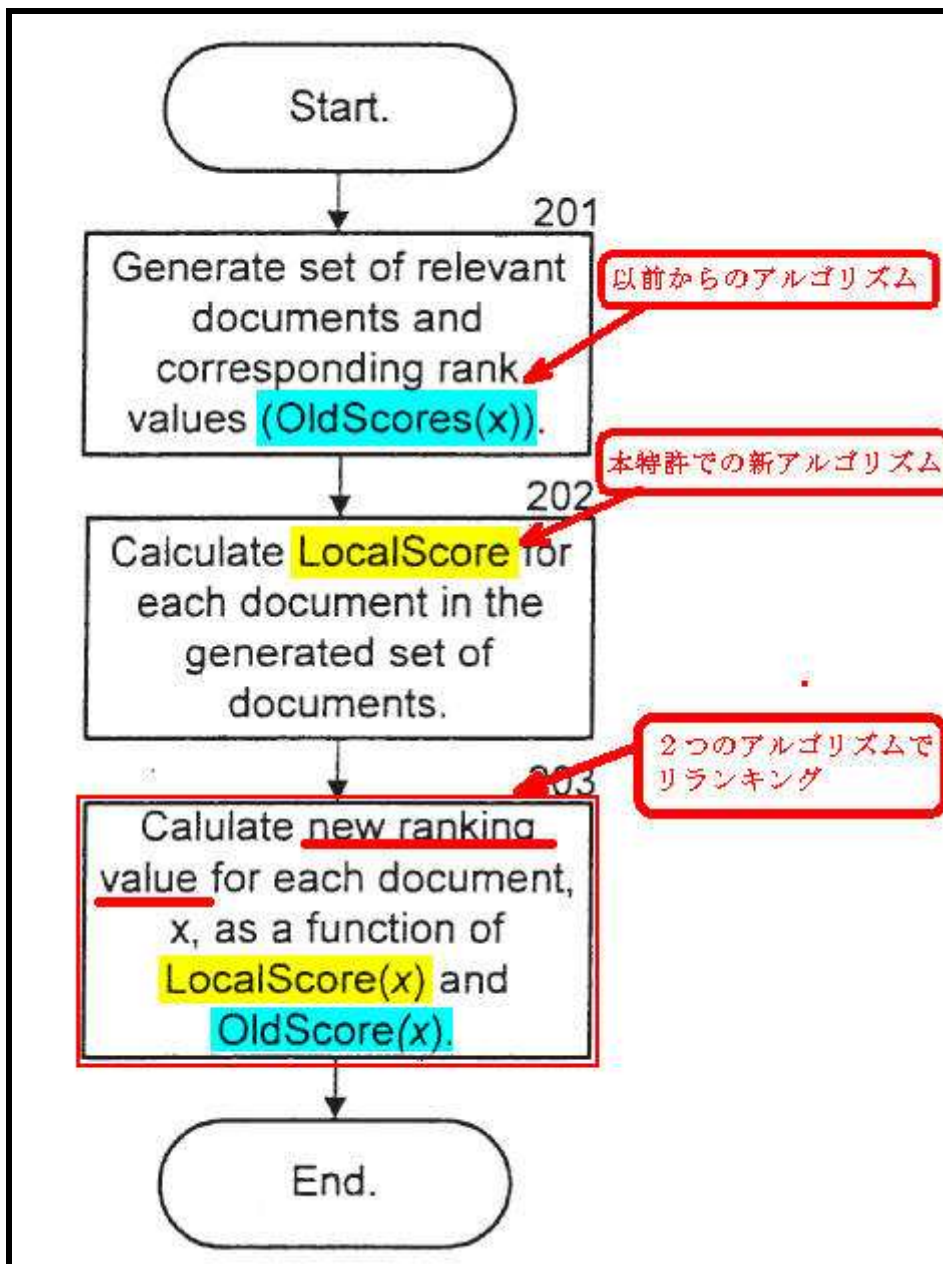


図1

この図1でGoogleが示しているのは、

1. **OldScore**と呼ぶ以前からのアルゴリズムに加えて新しいアルゴリズム、**LocalScore**(以下、**NEWアルゴリズム**)を導入する。そして、これの2つのアルゴリズムから検索順位を決定すること。
2. NEWアルゴリズムが、Googleが本特許で新しく追加したアルゴリズムであること。

そして、NEWアルゴリズムとは、リンク元のIPアドレスをどのように分析し、どのようにスコアリングしていくのかというアルゴリズムである。

では実際に、NEWアルゴリズムを見ていこう。

次の図2を解説していこう。

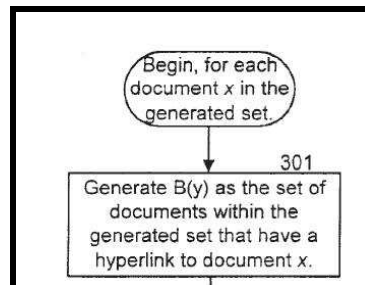


図2

まず最初に、この図2はDocument XというページにB(y)というリンク元があると述べている。

具体的な例を図3で説明しよう。

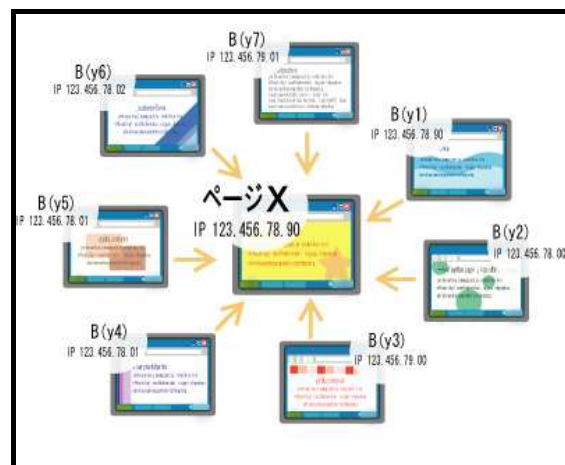


図3

DocumentX……リンクを得ているページX

ページXのリンク元… B(y1)~B(y7)

さらに、IPアドレスを次のように仮定する。

	IPアドレス
ページX	123.456.78.90
B(y1)	123.456.78.90
B(y2)	123.456.78.00
B(y3)	123.456.79.00
B(y4)	123.456.78.01
B(y5)	123.456.78.01
B(y6)	123.456.78.02
B(y7)	123.456.79.01

次に図4をご覧ください。

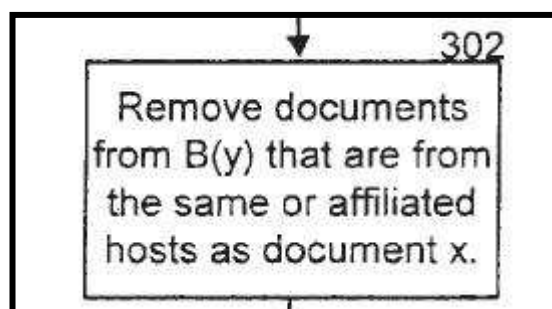


図4

ここで、新アルゴリズムはDocumentX(=ページX)と『同じ』か、もしくは『関連のある』ホスト(=IPアドレス)のB(y)のリンクスコアを削除すると記述している。

ここでいう、同じIPアドレスとは、具体的に言えば、クラスDまで同じIPアドレスのことである。

ここで、新アルゴリズムには次のルールが存在することがわかる。

NEWアルゴリズムの重要ルール:NO.1

ページXのIPアドレスとまったく同じ(=クラスDまで同じ)IPアドレスを持つ
リンク元B(y)のリンクスコア(リンクの価値)は削除される。

この重要ルールを図3の具体例で、考えてみよう。

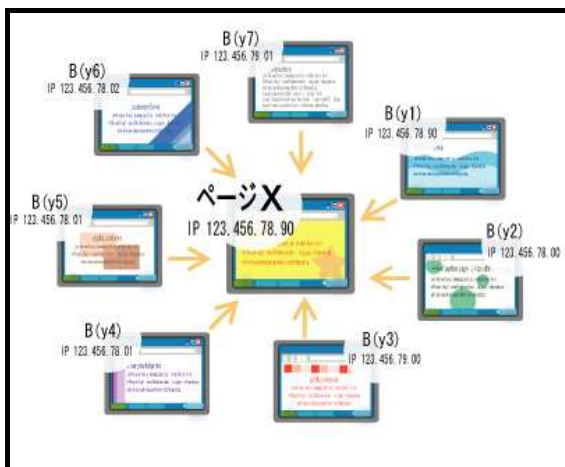


図3

	IPアドレス
ページX	123.456.78.90
B(y1)	123.456.78.90
B(y2)	123.456.78.00
B(y3)	123.456.79.00
B(y4)	123.456.78.01
B(y5)	123.456.78.01
B(y6)	123.456.78.02
B(y7)	123.456.79.01

ページXのIPアドレス = 123.456.78.90

クラスDまで同じIPアドレスを持つ、つまり全く同じIPアドレスを持つリンク元は:

B(y1) = 123.456.78.90

従って、B(y1)からのリンクスコア(リンクの価値)は削除され、B(y2)～B(y7)がリンク元としての価値が残る。

次に、『関連のある』ホスト(=IPアドレス)は、特許本文に次のように書かれている。

302). More specifically, let $IP3(x)$ denote the first three octets of the IP (Internet Protocol) address of document x (i.e., the IP subnet). If $IP3(x)=IP3(y)$, document y is removed from $B(y)$.

上記からも明らかなように、

『関連のある』ホスト(=IPアドレス)とはIP3のことを指している。

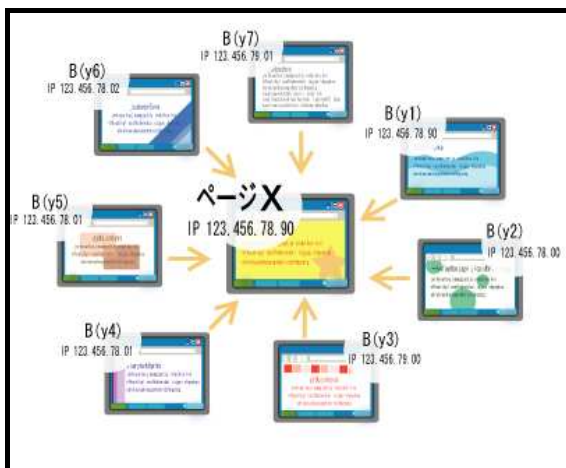
そして、IP3とはIPアドレスのクラスで言えばクラスCのレベルに該当する。

つまり、次のように定義される。

NEWアルゴリズムの重要ルール: NO.2

ページXのIPアドレスとクラスCレベルが同じIPアドレスを持つ
リンク元B(y)のリンクスコア(リンクの価値)は削除される。

この重要ルールを図3の具体例で、もう一度考えてみよう。



	IPアドレス
ページX	123.456.78.90
B(y1)	123.456.78.90
B(y2)	123.456.78.00
B(y3)	123.456.79.00
B(y4)	123.456.78.01
B(y5)	123.456.78.01
B(y6)	123.456.78.02
B(y7)	123.456.79.01

図 3

ページXのIPアドレス = 123.456.78.90

リンク元として残っているB(y2)~B(y7)を考慮する。

ページXとクラスCまで同じIPアドレスを持つリンク元は:

$B(y_2) = 123.456.78.00$

$B(y_4) = 123.456.78.01$

$B(y_5) = 123.456.78.01$

$B(y_6) = 123.456.78.02$

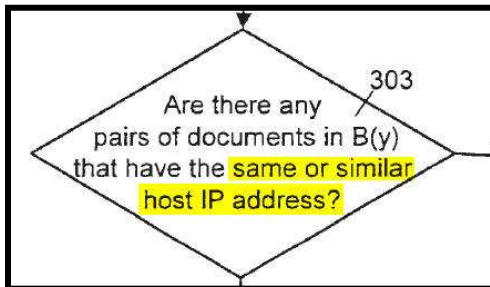
従って、4つのリンク元: $B(y_2)$, $B(y_4)$, $B(y_5)$, $B(y_6)$ からのリンクスコア(リンクの価値)は削除される。結果として、リンク元として価値があるものは、 $B(y_3)$, $B(y_7)$ だけとなる。

ここまでで、ページXのIPアドレスを元に、IPアドレスがリンクスコアにどのように影響するかについて解説してきた。

次に、ページXのリンク元 $B(y)$ 同士のIPアドレスをどのようにスコアリンクしているのかを説明しよう。

リンク元同士のIPアドレスにより算出されるリンクスコア

まず、特許では次のように説明されている。



B(y)同士で『同じ』か『似た』IP アドレスとは、前述と同じように、クラスDおよびクラスCが同じIPアドレスを指す。

さらに、次の内容

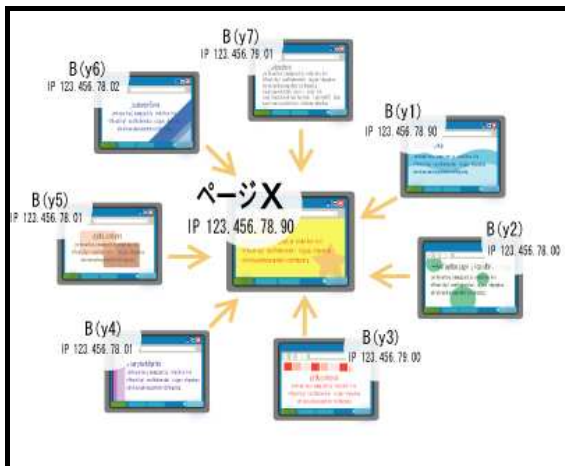
Re-ranking component 124 next compares all pairs of documents in B(y) for any pair in which IP3(first document of the pair)=IP3(second document of the pair), and removes the document of the pair from B(y) that has the lower OldScore value. (Acts 303–306). In other words, if there are

B(y)同士で同じIPアドレスや似た(クラスCレベル)で同じIPアドレスを持つ、B(y)を抽出する。そして、抽出されたB(y)の中からOldScoreアルゴリズムのリンクスコアで一番ポイントの高いB(y)のリンクスコア以外を削除する。

IP アドレス・アルゴリズムの重要ルール:NO.3

クラスC以下で同じIPアドレスを持つリンク元は、一番スコアが高いリンク元のリンクスコア以外は削除される。

この重要ルール2を図3の具体例で、また考えてみよう。



	IPアドレス
ページX	123.456.78.90
B(y1)	123.456.78.90
B(y2)	123.456.78.00
B(y3)	123.456.79.00
B(y4)	123.456.78.01
B(y5)	123.456.78.01
B(y6)	123.456.78.02
B(y7)	123.456.79.01

図3

ページXへリンクスコアを与えているページ(リンク元)は、前回のルール適用後、B(y3) , B(y7)である。

これら2つのリンク元IPアドレスは:

B(y3) = 123.456.79.00

B(y7) = 123.456.79.01

クラスC以下で同じIPアドレスである。

仮に、B(y3)のオールドスコア > B(y7)のオールドスコア とすると、

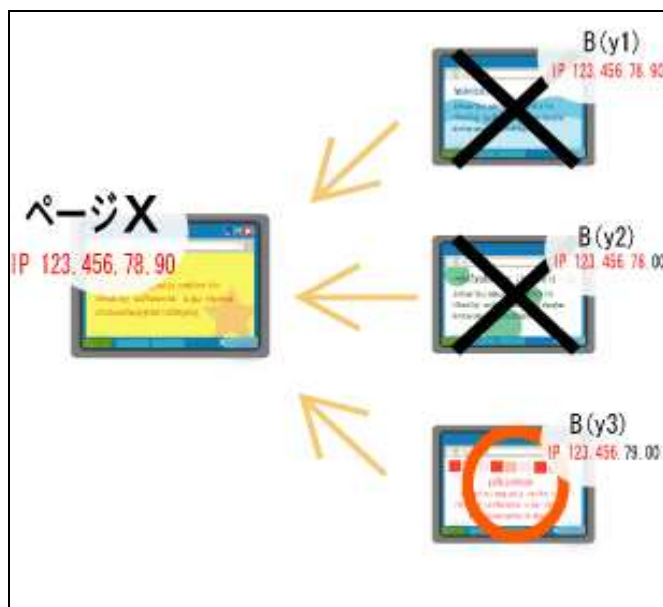
この重要ルール2の適用で、B(y7)のリンクスコアは削除され、B(y3)のリンクスコアのみカウントされる。

そして、最終的にリンク元B(y)の中で、クラスC以下で同一IPアドレスの場合、オールドスコアの一番高いものだけが残る。

こうして、B(y)の中で最終的に残ったページがオールドスコアの高い順番に並べられ、LocalScoreアルゴリズムに適用される。

それでは、もう一度、NEWアルゴリズムによって明かされるIPアドレスとリンク構造について考察しよう。

1. DocumentX(ページX)にリンクしているドキュメントを B(y)とする。
リンク元である B(y)の中で、DocumentX とクラスC以上でIPアドレスが同じページは B(y)から削除される。



2. 次に、リンク元 B(y)同士のIPアドレスを比較する。リンク元同士のIPアドレスがクラスC以上で同じページは、オールドスコアが一番高いもの以外全て削除される。

つまり、DocumentXへのリンク元となるページのIPアドレス(特にクラスCレベル)によって、検索順位で重要なリンクスコアに影響を及ぼすことが解明できる。

ここまでのポイントを整理してみよう。

IPアドレス分散特許からわかったこと

IPアドレスにより算出されるリンクスコア3つの重要ルール

ページXのIPアドレスとまったく同じ(=クラスDまで同じ)IPアドレスを持つリンク元B(y)のリンクスコア(リンクの価値)は削除される。

ページXのIPアドレスとクラスCレベルが同じIPアドレスを持つリンク元B(y)のリンクスコア(リンクの価値)は削除される。

クラスC以下で同じIPアドレスを持つリンク元は、一番スコアが高いリンク元のリンクスコア以外は削除される。

最後まで、お読みいただきありがとうございます。

今後も寡作ながら、アルゴリズムの特許を解明したレポートをメルマガで提供していきますので、ご期待ください。

ところで、本レポートを読んで、なかには、ここまでのことは、自社(自分)には不可能だ。

と、諦めの気分すら抱いた方もおられるかもしれません。

しかし一方、これはライバルに差をつける大きなチャンスだ！

と、感じた前向きな方も大勢おられると思います。

もし、あなたもそうであるなら、

あなたにピッタリのレンタルサーバーがあります。

月々290円からクラスC以上でIPアドレスを5~125まで分散できるレンタルサーバーです。

詳しいサービス内容は下記をクリックしてください。

IPアドレス分散レンタルサーバー

詳しい情報は >> <http://www.123server.jp/>