

## 深い階層情報に対応する円状配置を用いた アクセスログの詳細度制御付き可視化システム

細田 英幸 牧野 光則

中央大学大学院理工学研究科

### 1 序論

現在, WWW (World Wide Web) 上には数十億を超える Web ページが存在する. Web サイトは, 情報の発信や収集に多くの役割を果たし, ユーザは情報を得るために様々な Web サイトを訪れる. ユーザは訪れた Web サイトが有用でないと判断したら, すぐに他の Web サイトへと移ってしまう. そのため, Web サイト管理者は, アクセスログを解析し, 有益な情報に辿り着けるように適切にリンクすることが重要になる. アクセスログとは Web ページを訪れたときにサーバに蓄積されるデータのことであり, 解析することにより, ユーザがどのファイルに多くアクセスしたか, どのようなページの移動経路をとることが多いかがわかり, サイト構造の問題点を知ることができる.

膨大な情報量となるアクセスログからアクセス傾向を効率的に把握するために, 情報可視化技術が注目されている. アクセスログ可視化に利用できる代表的な情報可視化技術として, 木構造型の Cone Tree[1], 納豆ビュー [2], 画面分割型の Treemaps[3], 長方形入れ子構造型のデータ宝宝箱 [4] が提案されている. しかし, 木構造型にはフォルダ構造やページ移動経路を把握しやすいという長所があるが, 表示領域が多く必要かつアクセス数の比較もしくにくいという短所がある. また, 画面分割型と長方形入れ子構造型には表示領域が少なくアクセス数の比較がしやすいという長所があるが, ページ移動経路を同時に表現できないという短所がある. そこで本研究では, 木構造型をベースにしながら, 詳細度制御と長方形入れ子構造型の特徴を取り入れることにより, アクセス傾向を把握するために重要になる, アクセス数とフォルダ構造とユーザのページ移動経路を注目状況に応じた詳細度で同時に表現できるシステムを提案する.

## 2 提案手法

### 2.1 提案の概要

対象は Web サイト管理者とし, アクセス数, 深い階層構造を持つフォルダ構造, ならびにユーザのページ移動経路の三者の可視化を目的とする.

ファイルはアイコンで, アクセス数は基準面からの高さで, ページ移動経路はファイル間を結ぶ直線で表現する. フォルダ構造は初めに概要を表示し, 対話的操作により上位階層に視点が接近した場合に詳細を表示する. これにより, 注目状況に応じた詳細度で三者の同時表現ができる.

### 2.2 描画方法

$xz$  平面が初期状態で水平,  $y$  軸が画面上方, 初期注目階層は第 1 階層, 第 1 階層の座標は  $(0, 0, 0)$  とする.

#### 2.2.1 ファイルの表現方法

アクセスされたファイルは球体のアイコンで表現する. アクセスされていないファイルはアクセスログに記録されないため, 無視する. また, 注目階層にファイルが存在する場合, ファイル名とアクセス数を表示する.

#### 2.2.2 アクセス数の表現方法

アクセス数は  $xz$  平面からのアイコンの高さで表現する. 位置が高いアイコンほどアクセス数が多いことを表す. アイコンの高さは, 以下の式で求める.

$$\text{高さ} = \frac{H \times \text{ファイル毎のアクセス数}}{\text{全ファイル中の最大アクセス数}} \quad (1)$$

ここで,  $H$  はユーザが任意に設定できる高さの最大値を表す.

#### 2.2.3 フォルダ構造の表現方法

注目階層と注目階層から 1 階層下位のフォルダはファイルとフォルダを  $xz$  平面上に円状に配置することで表現する. また, 注目階層から 2 階層下位のフォルダは, 初期状態ではファイルとは異なる色の球体のアイコンで表現するとともに, ファイル数とフォルダ数を表記する. 内側に配置する注目階層から 1 階層下位の  $n$  個のファイルとフォルダの座標  $(X_{i_{in_j}}, Z_{i_{in_j}})$  ( $j = 0, \dots, n-1$ ) と, 外側に配置する注目階層のファイルとフォルダの座標  $(X_{i_{out}}, Z_{i_{out}})$  は以下の式で求める.

$$\begin{aligned} X_{i_{in_j}, out} &= x_{in_j, out} + R_{in_j, out} \cos(2\pi i_{in_j, out} / m_{in_j, out}) \\ Z_{i_{in_j}, out} &= z_{in_j, out} + R_{in_j, out} \sin(2\pi i_{in_j, out} / m_{in_j, out}) \\ &\quad (i_{in_j, out} = 0, \dots, m_{in_j, out} - 1) \end{aligned} \quad (2)$$

ここで,  $x_{in_j, out}$  はフォルダの座標,  $m_{in_j, out}$  はファイルとフォルダの合計数を表す. また,  $R_{in_j, out}$  はフォルダの半径を表し,  $R_{in_j}$  は球体であるアイコンの半径  $r_{icon}$  を基準に,  $R_{out}$  は各  $R_{in_j}$  の中で最大の半径  $r_{in_{max}}$  を基準にして, 隣り合うアイコンが重ならないよう以下の式で決定する.

$$\sqrt{(R_{in_j, out} \cos \theta_{in_j, out} - R_{in_j, out})^2 + (R_{in_j, out} \sin \theta_{in_j, out})^2} = 2r_{icon, R_{in_{max}}} + L_{in, out} \quad (3)$$

ここで、 $L_{in,out}$  はユーザが任意に設定できるアイコン間の余白の長さを表し、 $\theta_{in,out}$  は隣り合うアイコンがなす角度を表す。

注目階層から3階層より下位のファイルとフォルダは初期状態では画面に表示せず、対話的操作により視点  $(V_x, V_y, V_z)$  と1階層上位の各フォルダの中心位置  $(P_{x_j}, P_{y_j}, P_{z_j})$  の距離が以下の式を満たした時、注目階層を移して詳細を表示する。

$$\sqrt{(V_x - P_{x_j})^2 + (V_y - P_{y_j})^2 + (V_z - P_{z_j})^2} = R_{in_j} + D \quad (4)$$

ここで、 $D$  はユーザが任意に設定できる距離を表す。また、詳細表示後、対話的操作により注目階層から離れて、視点と注目階層の距離がユーザが任意に設定できる距離  $D_{back}$  になった時、注目階層を戻して概略表示に戻る。

#### 2.2.4 ページ移動経路の表現方法

ユーザのページ移動経路はアイコンを結ぶ直線で表す。また、全ての移動経路の中でユーザが一番多く取った経路の移動数に対する割合が多いほど、直線の色を濃くする。

### 2.3 対話的な処理一覧

対話的操作方法を、表1に示す。

表1: マウス操作

操作	機能
左ドラッグで上下左右に移動	視点を回転
右ドラッグで上下に移動	視点を前後に移動
視点をフォルダに近づける	詳細表示
視点をフォルダから遠ざける	概略表示

## 3 シミュレーション

WWW上に約800ページを公開しているWebサイトの1日のアクセスログを提案手法により可視化した結果、及びシステムの実行結果を図1、図2に示す。このWebサイトは4階層までのフォルダがある。図1は概観表示画面、図2は視点移動により詳細表示した画面である。

可視化結果から、このサイトは第2階層のフォルダにファイルが多いことがわかる。次に、サイト内のリンクのほとんどがフォルダ内で完結していることから、関連しているファイルが適切にフォルダにまとめられていることがわかる。

また、フォルダ構造を平面に展開し、アクセス数を高さで表現したことにより、従来の可視化手法では実現できていないアクセス数とフォルダ構造とページ移動経路を同時に把握できるようになっている。

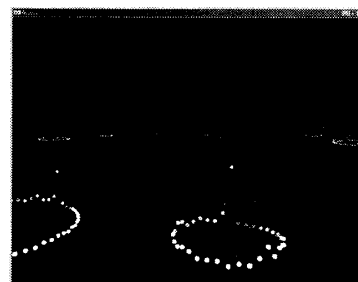


図1: 実行例1

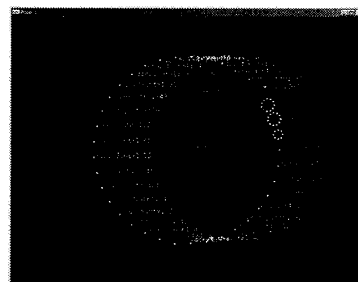


図2: 実行例2

## 4 結論

本研究では、Webサイト管理者がアクセス傾向を効率的に把握するためのアクセスログ可視化システムを作成した。木構造型をベースにした新たな可視化手法を提案することにより、深い階層情報を持つ場合でも、アクセス数とフォルダ構造とユーザのページ移動経路を注目状況に応じた詳細度で同時に表現できるシステムを実現した。

## 謝辞

本研究の一部は、文部科学省科学研究費補助金、(財)テレコムエンジニアリングセンター公益的調査研究、中央大学理工学研究所共同研究の援助をそれぞれ受けた。

## 参考文献

- [1] G. G. Robertson, J. D. Mackinlay, and S. K. Card, "Cone Trees: Animated 3D visualizations of hierarchical information", Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'91), pp. 189-194. ACM Press, 1991.
- [2] 塩澤秀和, 西山晴彦, 松下温, "「納豆ビュー」の対話的な情報視覚化における位置づけ", 情報処理学会論文誌, vol38, No. 11, Nov. 1997, pp2331-2342
- [3] Brian Johnson and Ben Shneiderman, "Treemaps: A space-filling approach to the visualization of hierarchical information structures", Proc. of the 2nd International IEEE Visualization Conference, pp. 284-291, San Diego, October 1991.
- [4] 山口, 池端, 伊藤, 梶永, データ宝石箱を用いたウェブアクセスログの視覚化, 第30回可視化情報シンポジウム, 2002.