

# ランキング機構を備えた 小規模サーチエンジンシステム

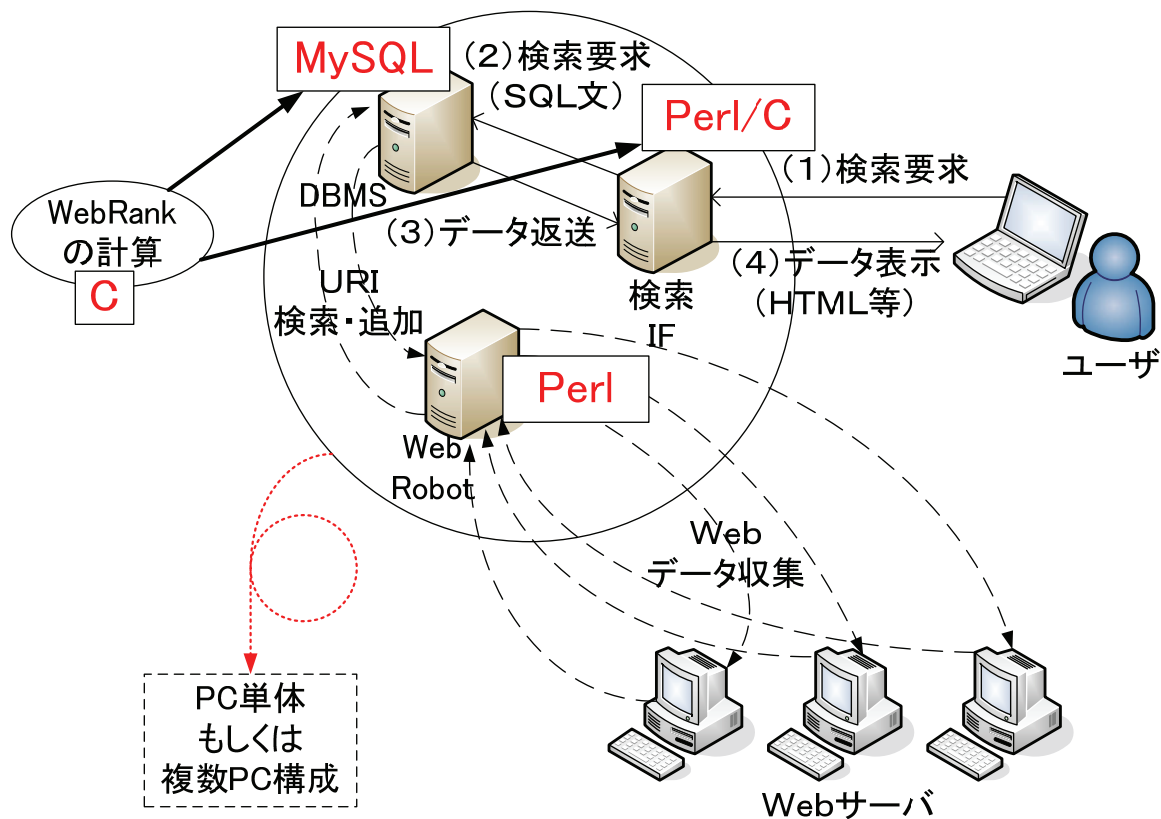
静岡理科大学 総合情報学部 幸谷 智紀

<http://na-inet.jp/>

(資)わいにじ 竹口 友大

<http://www.y2ji.co.jp/>

## 1. システム概要



普通のパソコン単体で動作する小規模サーチエンジンシステムです。これは次の4つのソフトウェアで構成されています。

MySQL・・・DBMS(データベース管理システム)

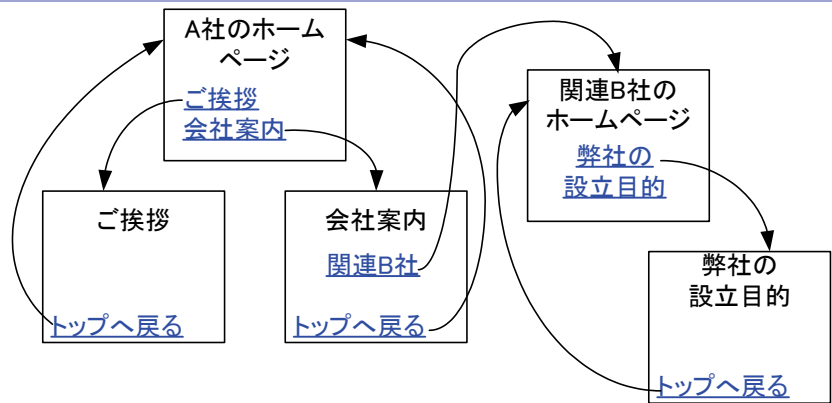
Web Robot・・・Webデータを自動収集するPerlスクリプト

検索IF・・・キーワード検索が可能なWebインターフェース

WebRank計算・・・リンク構造を元にしたランキングの計算を実行

## 2. ランキングの計算

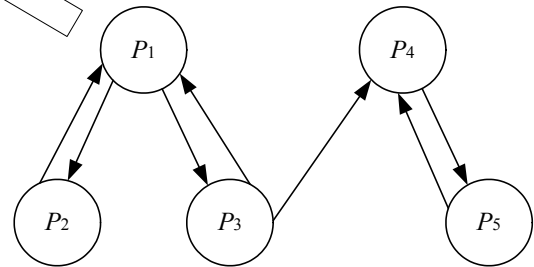
ホームページ間のリンクの構造に基づいて、ランキングを計算するための行列を導出します。



$$\begin{bmatrix} r(P_1) \\ r(P_2) \\ r(P_3) \\ r(P_4) \\ r(P_5) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r(P_1) \\ r(P_2) \\ r(P_3) \\ r(P_4) \\ r(P_5) \end{bmatrix}$$

リンク構造に基づいて  
ランキングの計算を行う

リンク構造を  
取り出す



$r(P_i) \cdots P_i$ のランキング

$$\Rightarrow \mathbf{r} = \mathbf{Hr}$$

$$\Rightarrow \mathbf{r} = \left[ \alpha \mathbf{H} + \frac{1}{n} \mathbf{a} \{ \alpha \mathbf{e}_i + (1 - \alpha) \mathbf{a}^T \} \right] \mathbf{r}$$

を  $\mathbf{r}$  について解く！

ランキング計算は我々オリジナルのソフトウェアで高速に実行します。

## 3. 検索の例

「大学」で検索

計算されたランキン  
グ値の大きい順に  
検索結果が並ぶ。