

# 自己組織化マップ (SOM) を用いた OSS 開発支援コミュニティの 提案と評価

河村 進吾<sup>†</sup> 田路 真也<sup>†</sup> 森 秀樹<sup>†</sup> 上原 稔<sup>†</sup>

東洋大学大学院工学研究科情報システム専攻<sup>†</sup>

## 1. はじめに

オープンソース・ソフトウェア (OSS) の開発は、集合知 (Wisdom of crowds) の性質を持つ代表例である。より質の高い集合知を発揮するためには、開発経験のないライトユーザの意見や発想の取り込みが不可欠である。筆者らが提案するシステムは、集合知の持つ性質 (多様性、独立性、分散性、集約性) を出発点にしたユーザ参加型開発支援 Web コミュニティである。ここでは、開発経験のないライトユーザを開発参加者として取り込み、今までのソフトウェア開発モデルでは得られなかった新しいアイデア、人材を発掘することを目指す。本稿では自己組織化マップ (SOM) を用いて、開発グループ編成時における多様性を保障する新しいシステムを提案する。

## 2. 自己組織化マップ (SOM)

SOM は、ヘルシンキ大学の Kohonen[1]によって考案された、教師無し学習型ニューラルネットワークの一種である。ヒトの脳における視覚野をモデル化した、入力層と出力 (マップ) 層の 2 層からなる多層式ネットワークを持つ。SOM のアーキテクチャを図 1(a)に示す。

$$d_i = \sqrt{\sum_{j=1}^n (x_j - \omega_{ij})^2} \quad (1)$$

$x$ : 入力ベクトル,  $\omega$ : 参照ベクトル

$$\text{if } i \notin N_c, \text{ then } h_{ci} = \alpha_0(1-t/T) \quad (2)$$

$$\text{else } h_{ci}(t) = 0$$

$$N_c(t) = N_c(0)(1-t/T) \quad (3)$$

SOM では、多次元の入力を「入力ベクトル ( $X = x_1, x_2, \dots, x_n$ )」として与え、2次元格子状に配置されたニューロン (ノード) との間の距離 (ユークリッド距離: 式 1) を計算する。ノードは初期値において重み値を持っており、重み値と入力値の間で、最も距離が近いノードを勝者ノードとして選択し、そのノードと近傍の重みを近傍関数と呼ばれる評価関数 (式 2, 3) を用いて繰り返し更新することによって、初期状態においてランダムな重み値を持っていた各ニューロンが、対象の特徴を学習する。

このような特徴から、SOM はクラスタ分析やデータマイニングの手法として注目されている。

SOM によって得られたマップから、マップ上で近い距離のユーザ同士ほど類似性をもったユーザであることがわかる。本研究では、提案する Web コミュニティにおいて、参加者の多様性を評価する手段として SOM を選択した。

## 3. オープンソースコミュニティ C-DOS

C-DOS は、筆者らが提案する OSS 開発のためのユーザ参加型開発支援 Web コミュニティである。C-DOS では、プロジェクト提案コミュニティと、プロジェクト開発/進行コミュニティの 2 つの性格を有している。

## 4. 参加メンバーの特徴分析

C-DOS では、希望者は条件を問わずにコミュニティに参加登録をすることができる。ユーザ登録の際に、ユーザは自分のソフトウェア開発スキルについての自己評価を行なう。登録が完了したユーザは、開発候補者 DB に登録されるが、同時にコミュニティへの参加者としての権利も与えられる。ソフトウェア開発のアイデアがある、もしくは既存のソフトウェアに対する改良点、要望などがあれば、すぐにプロジェクトとして話題を提供することができる。ユーザ発のプロジェクト提案は、SOM を通して開発者候補として認められたコミュニティ参加者に招待メッセージが送られ、参加者を募る。参加希望者が一定数を超え、この提案に対して、コミュニティ内で一定の賛同を得られた場合、プロジェクトは正式に発足する。だが、発案者は「発案するだけ」でプロジェクトのリーダーにはならない。

OSS ではアジャイル開発モデルが頻繁に使われている。これは、要求や仕様の変更に柔軟に対応できるためであるが、このモデルを有効に利用することができるのは、限られた有能なプログラマに限られる。そこで、筆者らはスパイラルモデルを基に、C-DOS におけるプロジェクトの回転の速さに対応するために、図 1 のようなモデルを考案した。

A Method of Project Member Arrangement of using SOM for a OSS Support Community

<sup>†</sup> Department Open Information Systems, Graduate School of Engineering, Toyo University

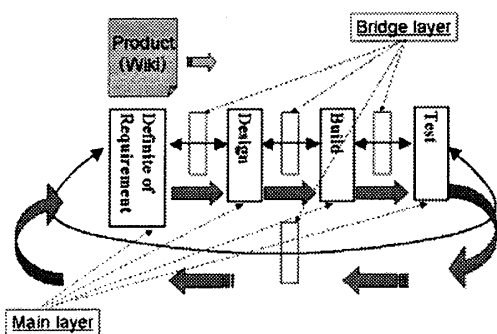


図 1. C-DOS の開発モデル

各工程（レイヤー）は要求定義，設計，実装，テストの4つを基本とし，プロジェクトが認証された後，プロジェクト参加者らは開発メンバとして，SOM 分析の結果，各レイヤーへ配属される．各レイヤーでの仕事が完成し次第，隣のレイヤーにその仕事を託すことになる．また，各レイヤー間にサブ工程として，ブリッジレイヤーを設ける．ブリッジレイヤーに配属されたメンバは両隣のレイヤーへの参加資格を持つ．ブリッジレイヤーを置くことで，プロジェクトの一貫性・整合性の維持と，前レイヤーの結果を受けて，速やかな次レイヤーへの橋渡しを目指す．ブリッジレイヤーのメンバもまた，他のレイヤーと同様にシステムによって割り振られるが，ブリッジレイヤーの効果の強さは，そのレイヤーに配置される人数の加減で調整する．

### 5. 実験

コミュニティ参加希望者の興味・趣味と，ソフトウェア開発に関する知識について SOM 分析を行なう．ソフトウェア開発に関する知識については，SWEBOK[2]を用いて評価する．これは，大学卒業後に少なくとも4年間の実務経験を経たプロフェッショナルが身に付けるべき知識の体系として，IEEE を中心に2001年に制定されたものである．なお，実験環境としてオープンソース数理解析環境の R を用いた．実験状況を表1に示す．

表 1. SOM 分析の概要

入力データ	マップの形状	学習回数
26次元×20件	3×3 四角形	10000

### 6. 結果

上記に挙げた入力フォーマットに沿った20人分のダミーデータを用いて SOM 分析を行った．ID：#1～#10は開発経験がある，#11～#20までは開発経

験がないユーザのデータを設定した．図2に出力結果を示す．得られたマップから，開発経験のないユーザが同じユニットに属し，逆に開発経験のあるユーザは離れたユニットに属していることがわかる．SOM 分析を行なった場合，ノードの重み値はランダムに初期化されるため，分析ごとに異なったマップ配置を得ることがある．だが，いずれの場合でも開発経験を用いてユーザの特徴を分類させることができた．

□ : No development experience    ■ : Development experience

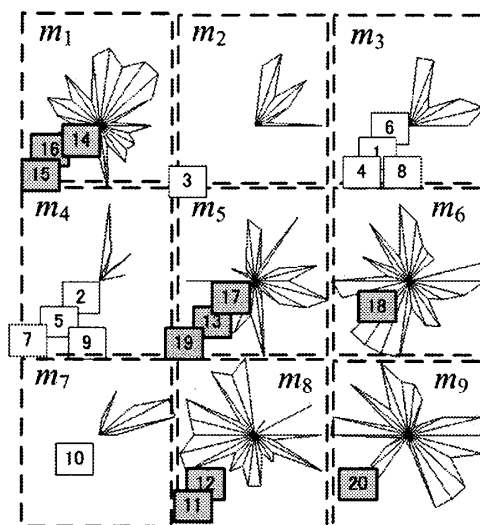


図 2. SOM 出力結果

### 7. むすび

本稿では，OSS 開発支援 Web コミュニティ「C-DOS」における，SOM を用いたプロジェクトメンバの分類法について考察した．スキルに関するマップのみでなく，個人の性格や，心理テストのアンケートや調査の結果からもマップを生成することで，さらに詳しい分類のための手がかりになることが期待できる．今後は，システムの実装に向けたスケーラビリティ，実現可能性の検証と，SOM 分析のための詳細なパラメータの検討，およびシステム全体の評価方法の策定が当面の課題である．

### 参考文献

- [1] T. Kohonen: 著，Self-Organizing Maps. Springer, New York, 1995.
- [2] SWEBOK: IEEE Computer Society (<http://www.swebok.org/>)