

## 商取引ネットワークにおける不審ユーザ判定モデルの評価と実装

小林 真雄†

伊藤 孝行‡

†名古屋工業大学情報工学科

‡名古屋工業大学大学院大学産業戦略工学専攻

### 1 はじめに

昨今、インターネット上における商取引が盛んであり、オンラインショッピングやネットオークションはその代表的なものである。特にインターネットオークションの利用者は年々増加している。しかし、近年、詐欺などの不正行為がメディアで報じられている。日本最大手 Yahoo!オークションは多くの利用者を抱える一方、多くの詐欺事件も発生している。サイト側はオークションの監視、データの解析を強化することで不審者を排除する働きかけを行っている。しかし、不正行為を防ぎきれていないのが現状である [1]。そこで、本研究では、オークションの取引ネットワーク解析システムを実装している [2]。本論文では、オークション上のユーザデータから不審な履歴を持つユーザを検出するためのユーザモデルについて論述する。

### 2 不審者判定機能

本システムはオークションユーザデータを様々な視点から解析、提示するシステムであり、不審者判定機能を持つ。本機能では一般的なユーザをモデル化することで、一般的なユーザモデルから逸脱するユーザを検出する。本論文での提案手法は、不審者判定における指標にオークションの利用法に基づく数値を採用する。本システムは各ユーザのオークションでの取引における立場の違いを AucType と呼ぶ。取引における立場とは、売り手または買い手のことである。ユーザの各取引での出品率、つまり売り手側である比率を AucType という数値で表している。不審者の判定機能は、各オークションユーザの持つ属性値の母体に対する偏差値を用いて判定する。

### 3 システムの実装

今回の実装は次の 2 つの要素からなる。一つは解析に用いるデータを収集するためのサイト巡回プログラムの実装である。もう一つは先に述べたようなユーザデー

タを解析、提示するシステムである。共に JDK1.5 で実装し、データベースとして MySQL5, MacOSX 10.5.1 上で運用している。また、本システムでのオークションデータは Yahoo!オークションに準拠している。

まずデータ収集について述べる。本研究では解析の元となるデータが無い場合、実際のオークションサイトから取得してくる必要がある。そこで半自動的オークションサイトを巡回しデータベースにデータを蓄積させる巡回ロボット、クローラーを作成した。クローラーに取得させる 1 ユーザあたりのパラメータを表 1 に示す。

表 1: FraudParameter の構成

Field	説明
ID	ユーザのオークション ID
time	データの取得時間
own_rep	解析対象ユーザの評価値
partner_ave	対象ユーザの取引相手の平均評価値
auc_type	対象ユーザのオークションタイプ (出品率)
bad_ratio	対象ユーザに対する評価の悪評価比
banned_num	対象ユーザの取引相手中の登録削除者、停止者数
banned_ratio	対象ユーザの取引相手中の登録削除者、停止者比
new_num	対象ユーザの取引相手中の新規者数
new_ratio	対象ユーザの取引相手中の新規者比
label	解析対象ユーザが「ストア」であるか
span	取引履歴から抽出したオークション利用期間

次にデータ解析について述べる。クローラーが収集したデータを解析し、その結果を効果的に表示するために本研究では Java ライブラリである JFreeChart を用いたシステムを作成した。本システムでは、AucType を中心としたデータ解析が行われ、円グラフや散布図など各種表示形式で多角的にオークションデータを示す。図 1 は本システムの出力画面例である。左上図は収集したユーザデータを AucType で分類した円グラフを示す。右上図は AucType で分類されたグループ毎の平均評価値の棒グラフを示す。左下図は収集した全ユーザを AucType と悪評価比でプロットした散布図を示す。右下図は個々のユーザを全ユーザを母体としたときの偏差値を表す九角形レーダーグラフを示す。図 1 で示すもの以外にも本システムでは個々の要素についての表示機能を持つ。

L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub> Style File for National Convention of IPSJ

†Masao KOBAYASHI ‡Takayuki ITO

†Department of Computer Science and Engineering, Nagoya Institute of Technology

‡Master course of Techno-Business Administration, Nagoya Institute of Technology

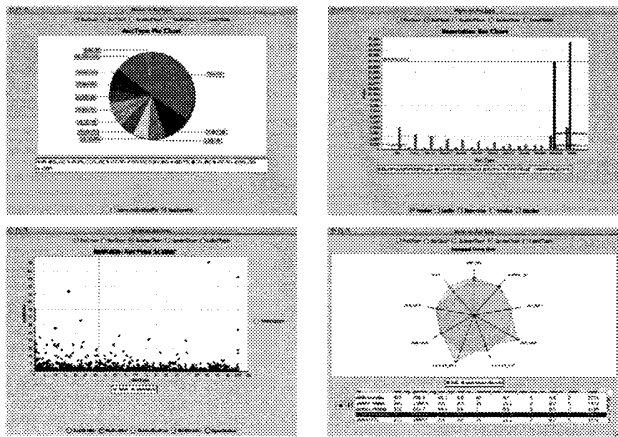


図 1: 解析システムの出力画面

#### 4 実験と評価

本システムの評価として、不審者判定機能の計算母体別の判定能力の比較を行う。実験設定として、実験に用いた総データ数は 16403(個人ユーザのみ)。判定条件には、『悪評価が高く、且つ評価値が異常な』ユーザを不審者とした。数値的には先に示したユーザパラメータの own rep と bad ratio を変化させ、システムが不審者とするヒット数の推移を検証する。図 2, 及び図 3 では、X 軸に閾値 1 として各ユーザの持つ各々の母体での own rep の偏差値の絶対値をとる。Y 軸に閾値 2 として悪評価比の偏差値を表す。実験での平均値を偏差値の 0 としている。また、Z 軸にヒット数をとる。ヒット数とは閾値 1, 及び 2 を共に超える偏差値を持つユーザの総数である。

評価として、全データを母体とした時のヒット数と、AucType 別を母体とした時のヒット数の傾向の違いについて示す。図 2 では、閾値 1 が 2~3 で急激にヒット数が減少している。これは全ユーザを母体とした時の各ユーザの評価値の閾値が 2~3 に集中していることが分かる。一方、図 3 では閾値 1 の変化による急激なヒット数の変化は起こっていない。これは偏差値計算におけるデータの母体に AucType による分類を用いることで、同じようなオークションの使い方をしている人の中ではそれほど逸脱したユーザはいないことを示す。以上から、全データを母体とした時には不審者としてのヒット数に評価値(閾値 1)が大きく依存し、AucType 別に母体をとった時にはヒット数が閾値と反比例関係にあるのが分かる。このことから、不審者判定が同じ条件であっても全データを母体にとることで検出されない事例もある。

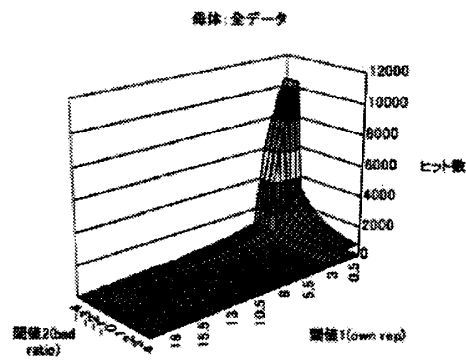


図 2: 母体を全データとしたときの不審者判定結果

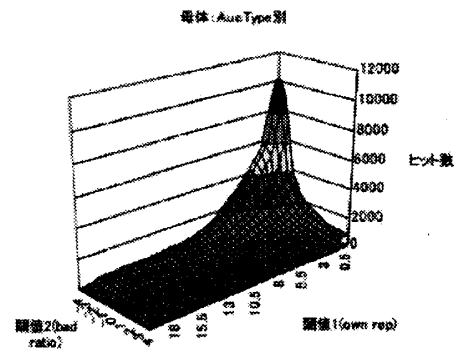


図 3: 母体を AucType 別でとったときの不審者判定結果

#### 5 まとめ

本論文ではインターネットオークションのユーザデータから不審者を判定するモデルの実装、評価について述べた。また、不審者判定機能として一般ユーザモデルからの逸脱者を判定した。実験において AucType を考慮したユーザモデルの有効性について述べた。今後の課題としては、更なる大規模データ取得と取得データを新たなアプローチで解析することによる各要素間の依存関係の発見が考えられる。また、詐欺者同定の観点からユーザデータに基づくアプローチの他に、ユーザ間の繋がりを利用する手法がある。

#### 参考文献

- [1] 八代 峰樹, "Yahoo!オークションにおけるユーザの悪用問題と対策," IEICE Technical Report AI2007-25(2007-12)
- [2] M. Kobayashi and T. Ito, "A Transactional Relationship Visualization System In Internet Auctions," IAT07,2007