Needs-scan/panelデータセットにおいて、クラスター数の増加とともにAUCが向上し続ける挙動を考察するため、クラスター構造を調査する実験を行った。

2ページ目以降の結果は各クラスター数においてRSCCMM-CFを適用し、高いAUCとなった際の、「各ユーザーが分類されたクラスターの数のヒストグラム」である。なお階級幅は1として表示している。

各クラスター数でdeltaを種々変化させ、10回試行で最も高いaucとなった際の結果を採用している

結果から分析できる点：

1. すべてのクラスター数において、「1つのクラスターにしか所属していないユーザー」が一定数存在する。
2. クラスターが増加していくとグラフ右側の「最も多くのクラスターに所属しているユーザーの総数」が増加していく。
3. ２以外のユーザーの総数はクラスター数が増えてもあまり変わらず、割合としては減っていく

現状の考察：

クラスター数が多くなってもAUCが低下しない理由：

クラスター数が増えていっても過剰な分割は行われず、ほとんどのユーザーが多くのクラスターに所属し、クラスター構造が破壊されないと考えられる。

分析点の2のようなユーザーの数が増えていき、その他のユーザー数はあまり変わらないことから、分析点の2のようなユーザーがモデルユーザー、その他のユーザーがマイノリティー、ノイズユーザー（正確な表現かは怪しい）というような分け方になり中間発表での議論のような仮説が説明できる？

* C=20

グラフ, 棒グラフ

自動的に生成された説明

* C=40

グラフ, 棒グラフ, ヒストグラム

自動的に生成された説明

* C=60

グラフ, ヒストグラム

自動的に生成された説明

* C=80

グラフ, ヒストグラム

自動的に生成された説明

* C=100

グラフ, ヒストグラム

自動的に生成された説明