

論文紹介

高田 雄太

2017年11月7日

はじめに

- **本日は下記論文の内容をご紹介します。**

- G. Wang, X. Zhang, S. Tang, H. Zheng, and B. Y. Zhao, “Unsupervised Clickstream Clustering for User Behavior Analysis,” in *CHI 2016*.
- <http://www.cs.ucsb.edu/~ravenben/publications/pdf/clickstream-chi16.pdf>

- **概要**

- クリックデータ(clickstream) からウェブサイト上のユーザ挙動を分類する手法ならびに分類した挙動を可視化するシステムを実装。加えて、本システムのユーザスタディを実施。
- 実SNSで取得した clickstream traces(142M events) を使用して評価した結果、未知のユーザ挙動 (誹謗中傷、Sybilアカウント)を発見。また、可視化システムによりユーザ挙動の理解を促進させることを確認。

背景

• SNS等のインターネットサービスの発展

- 各ユーザ挙動の要因/動機づけを理解するのは、UI/UX向上に有用だが、非常に困難
 - e.g., LinkedIn は学生, 専門家, リクルータ等の多様なユーザがおり、それぞれ異なる使い方をしているが、プロフィールからそれを特定するのは困難
- 「システムがどのように使われているか」のユーザスタディは既にあるが[3,34]、下記の制限が存在
 - 詳細なユーザスタディは、スケールしないこと（コスト）
 - ユーザは、自身を異なるユーザカテゴリへ自己判断できないこと
 - 既知の問題/仮説に基づくユーザ調査であること（未知の行動は不明）

• データドリブンアプローチは上記を解決

- 多くのシステムで時系列ユーザ挙動データ(clickstream)を収集
- clickstream を使ったユーザ挙動特定システムの要件は下記
 - noisy clickstream datasets でもスケールして機能すること
 - 未知のユーザ挙動を捕捉できること
 - 検知した挙動を理解しやすい形で解釈できること

データセット

Dataset	Time	# of Users	# of Events
Whisper	Oct.13–Nov.26 2014	99,990	135,208,159
Renren-Normal	Mar.31–Apr.30 2011	5,998	5,856,941
Renren-Sybil	Feb.28–Apr.30 2011	9,994	1,008,031

Table 1. Clickstream datasets from Whisper and Renren.

• 中国SNS: Whisper & RenRen

- Whisper: 10Mユーザがいる匿名メッセージアプリ
 - ランダムサンプルされたユーザの clickstream + userID を取得
- RenRen: 223MユーザがいるFacebookに似たSNS
 - 違いは「足跡機能」, i.e., 中国版 mixi
 - Sybil アカウントデータもあり

Category	Event Type	Events		Initiated By User?
		# (K)	%	
Browsing	View whisper	52437	38	Yes
	View popular feed	16008	12	Yes
	View nearby feed	5354	4	Yes
	View latest feed	2346	2	Yes
	View other feed	196	1	Yes
Account	Login	16994	12	Yes
Posting	Heart whisper	2156	2	Yes
	Upload image	1325	1	Yes
	Create whisper	1308	1	Yes
Chatting	Being blocked in chat	3271	3	No
	Block user in chat	3271	3	Yes
	Start a chat	2238	2	Yes
Notification	Receive notification	9680	7	No
	Whisper recommendation	2530	2	No

Table 2. Event types in the Whisper dataset. # of click events are presented in thousands. Events that are <1% are omitted for brevity.

Category	Event Type	Sybil Events		Normal Events	
		# (K)	%	# (K)	%
Friending	Send request	417	41	16	0
	Accept invitation	20	2	13	0
	Invite from guide	16	2	0	0
Photo	Visit photo	242	24	4,432	76
	Visit album	25	2	330	6
Profile	Visit profiles	160	16	214	4
Sharing	Share content	27	3	258	4
Message	Send IM	20	2	99	2
Blog	Visit/reply blog	12	1	103	2
Notification	Check notification	8	1	136	2

Table 3. Event types in the Renren dataset. # of click events are presented in thousands for Normal and Sybil users. Events with <1% of clicks are omitted for brevity. All of the events are user-initiated events.

提案システム：教師なしユーザ挙動モデリング

- **ユーザごとの clickstream から挙動パターンを補足するシステム**

- **Similarity graph**

- Clickstream を [イベント|時間間隔] の sequence データとし、タイムウィンドウ中の k-gram の出現回数に基づき、Polar Distance でシーケンス間の類似度を算出
- ノードをユーザ、エッジをユーザ間の類似度とする類似度グラフを構築

- **Iterative feature pruning**

- Divisive Hierarchical Clustering[13] で similarity graph を区切って、大まかな挙動分類を実施
- 区切ったクラスタからカイ二乗で主要な特徴を取得し、その特徴を再帰的に削っていく(pruning)ことで、クラスタを説明可能な階層構造を生成
 - 親クラスタ（一般的な特徴） > 子クラスタ（詳細な特徴）
- クラスタ密度を表す modularity[4] に基づき、再帰処理を停止

可視化手法

- 「メジャーな挙動カテゴリは何か？」 「異なる挙動間の関係は何か？」を視覚的に理解できるように可視化
 - 親子クラスタは階層表示とし、大きさをユーザ数を表現
 - Whisper: 107クラスタ, RenRen: 108クラスタ
 - 各クラスタをクリックすると詳細情報を確認可能
 - 挙動パターン、パターン出現頻度、カイニ乗スコア等

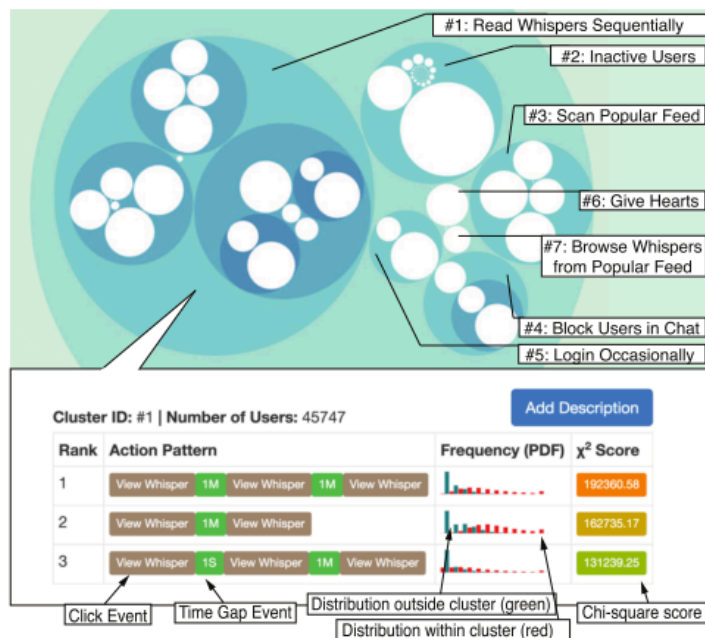


Figure 3. Whisper behavioral clusters. Cluster labels are manually input based on results of each cluster. The pop-up window shows users in Cluster #1 tend to sequentially read whispers.

多くのクラスタが
<5 の特徴で表現可能

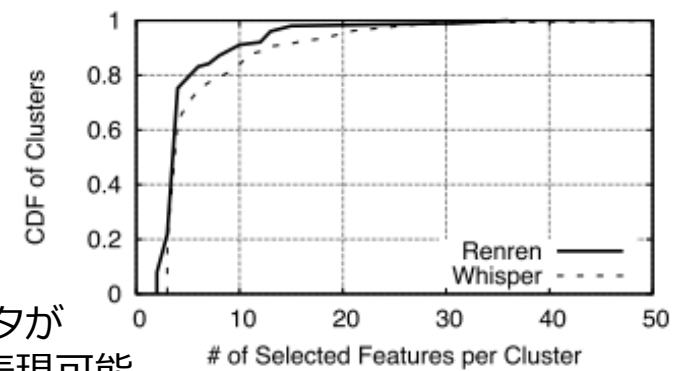


Figure 7. # of Selected features per cluster.

評価：ユーザスタディ

- **クラスタリング結果を用いて、「結果は理解しやすいか」「異なる解釈をしないか」を複数ユーザで評価**
 - ユーザ挙動の説明を一行で記述してもらおう、分からない場合は“N/A”
 - CS学部の学生 15名が参加で、参加者の時間節約のため、Whisper の代表的な階層クラスタのみテストに活用（意外とアッサリ）
 - 事前に10分間 Whisper アプリを使ってもらうとともに、事前に可視化ツールの使い方はレクチャー
- **ユーザスタディの結果**
 - 「結果は理解しやすいか」
 - 37クラスタから 530/555 (95.5%; 他は“N/A”) の説明文を収集
 - 平均 28.7min 以内にタスク完了 (46 sec/cluster)
 - 「異なる解釈をしないか」
 - CS学部の学生 3名(experts) を雇用し、記述内容の一貫性を確認した結果、60-80%の記述内容がおおよそ一致
 - 記述がバラバラだったクラスタは、階層が深い(より specific)傾向

評価：ケーススタディ

• Hostile Behaviors of Whisper Chatters

- private chat で他人をブロックしているアカウントの Cluster#4 を対象に、詳細に子クラスタを調査した結果、C#4のアカウントは活発である一方で、誹謗中傷の的になっていることが判明

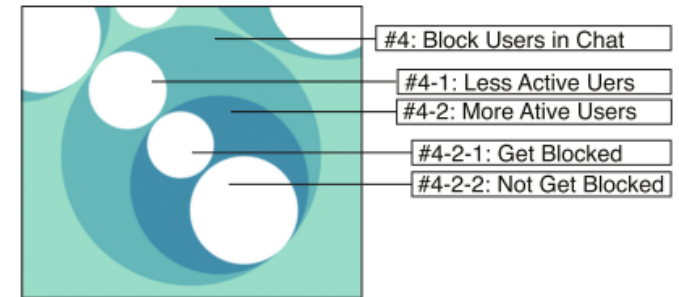


Figure 12. The sub-clusters within Cluster#4.

• RenRen Sybil Accounts

- 95%の Sybill account を正しく分類
- 詳細に子クラスタを見ていくと、大量のフレンド申請や“クローラ”のような働きをしており、ユーザの情報収集を実施している様子が浮き彫りになった

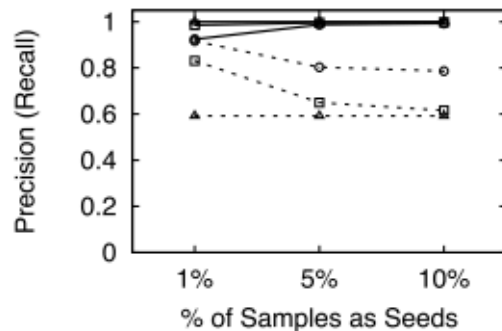
ID	Cluster Label	# of Users	FrdReq per Day	ProfileReq per Day	In/out FrdReq
S ₁	Friending in bulks	4064	25.13	0.30	0.002
S ₂	Friending quickly	1891	19.81	2.08	0.004
S ₃	Crawl profiles	1348	11.41	6.44	0.050
S ₄	Friending slowly	899	8.76	1.93	0.00004
S ₅	Receive FrdReq	129	25.65	3.43	0.286
#1	Normal users	6141	1.65	2.80	1.06

Table 7. Characteristics of users the 5 biggest Sybil clusters (S₁-S₅) and the normal user cluster. We add the cluster label based on the selected action patterns per cluster. “FrdReq” stands for “friend requests.”

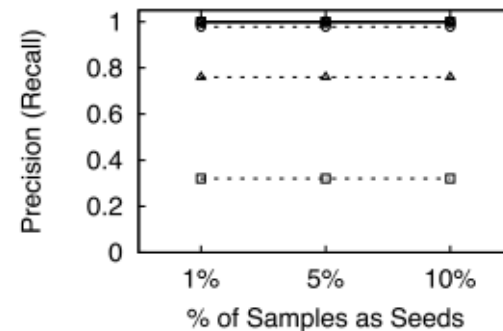
評価：クラスタ品質

• 既存クラスタリングアルゴリズムとの比較

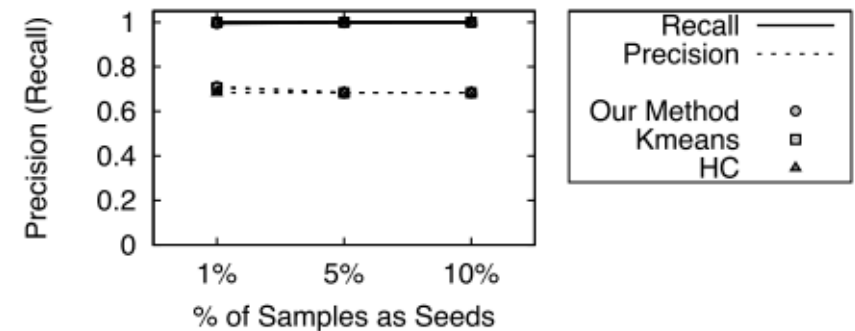
- 提案手法, K-means, Hierarchical Clustering を用いて、既知の Sybil account を正しく分類できるか精度を比較
- Ground truth の量 (1,5,10%) に対して、残りのアカウントを正しく分類できるかを計測
- 提案手法が最も正しく分類できていることを確認



(a) Sybils (Renren)



(b) Dormant Users (Whisper)



(c) Blocked Users (Whisper)

結論と今後の課題

- **複雑なユーザ挙動を補足するシステムを提案**
 - ユーザ挙動のクラスタを階層表示可能
 - 可視化により支配的なユーザ挙動やカテゴリを確認可能
 - 事前知識必要なく、未知のユーザ挙動も捕捉可能
 - e.g., Sybils, 行儀のよろしくないユーザ
- **異なるウェブサービスへの適用**
 - 例えば、Wikipedia, News, Q&A sites, or E-commerce における各イベントや特有イベントを捕捉したデータがあれば適用可能

所感

- **結果を様々な観点で評価している点が面白い**

- 「大量時系列データのクラスタリング」というよくある話題
- 「クラスタリングの結果を手動で確認したところ、概ね誤った分類はなかった」といった結果の記述がよく見られる
- 本研究ではユーザスタディ、ケーススタディ、既存手法との比較と様々なアプローチで評価していたため参考になった。

- **HCI な要素は少なめ？**

- “Mining Human Behaviors” セッション
- 「クラスタの説明文を人間に記述させ、その一貫性を人間に確認させる」
- データセット/システム/簡単なユーザスタディの説明方法は参考になる