

Study Group Workshop 2026 成果報告集

課題提供：富士通株式会社 富士通研究所

SGW2026

Weighed Signed Network における ランキングスコア設計

編集者)

浅尾 泰彦(九州大学マス・フォア・インダストリ研究所)

キム グンス(九州大学マス・フォア・インダストリ研究所)

池田 香凜(九州大学マス・フォア・インダストリ研究所)

この成果報告集について

➤ 成果報告集は複数の成果報告書をまとめたものです

- ✓ 成果報告書の下線付きの著者は、成果報告会(SGW5日目)で登壇したことを表します
- ✓ 登壇の有無に関わらず、成果報告書を成果報告集に採用することがあります

➤ 著作権および利用条件について

- ✓ 含まれる成果物の著作権は、原則として各成果物の創出者に帰属します
- ✓ 閲覧および閲覧・参照を目的としたダウンロードは自由です
- ✓ これらを超える転載、複製、改変、再配布、公衆送信等の自由利用の直接の対象は、原則としてSGW構成員です
 - 第三者による利用は、著作権法上認められる場合、または権利者・SGW構成員等の許諾がある場合に限られます
 - 製品開発等の商業的利用は、自由利用の対象外です
- ✓ 記載されたSGW2026由来のアイデアやノウハウそのものは、第三者も無償で自由に利用できます
- ✓ 詳細は以下のSGW規約をご確認ください
 - [国立大学法人九州大学マス・フォア・インダストリ研究所 スタディグループ・ワークショップ規約](#)

© 2026 Respective authors / creators.

Contents

➤ 課題提供者、モデレーター等

➤ 課題の説明

➤ 評価アルゴリズムの構築

- ✓ 藤江 勇仁 (九州大学大学院マス・フォア・イノベーション関係学府)
- ✓ 内海 忍 (理化学研究所数理創造研究センター)
- ✓ 門脇 大我 (九州大学大学院マス・フォア・イノベーション関係学府)
- ✓ 菅 陽佳 (九州大学大学院マス・フォア・イノベーション関係学府)
- ✓ 西 皇龍 (九州大学大学院数理学府)

➤ スコアリング手法の公理化について

- ✓ 岩倉 康樹 (九州大学大学院マス・フォア・イノベーション関係学府)
- ✓ 薄井 利基 (九州大学大学院マス・フォア・イノベーション関係学府)
- ✓ 杉林 直樹 (九州大学大学院数理学府)

➤ 時系列データに対する新しいfairnessとgoodnessの構築

- ✓ 寺田 怜央 (九州大学大学院マス・フォア・イノベーション関係学府)

Contents

- 時刻情報と評価件数情報を用いたスコアリング手法の修正
 - ✓ 崎原 雄大 (九州大学大学院マス・フォア・イノベーション関係学府)
 - ✓ 藤井 秀一
 - ✓ 力久 颯太 (九州大学大学院数理学府)

課題提供者、モデレーター等

▶ 課題提供者

- ✓ 坂巻 慶行(富士通株式会社富士通研究所)

▶ コーディネーター

- ✓ 徳田 悟(九州大学マス・フォア・インダストリ研究所)
- ✓ 白井 朋之(九州大学マス・フォア・インダストリ研究所)
- ✓ 北川 宜稔(九州大学 マス・フォア・インダストリ研究所)

▶ モデレーター

- ✓ 浅尾 泰彦(九州大学マス・フォア・インダストリ研究所)
- ✓ キム グンス(九州大学マス・フォア・インダストリ研究所)
- ✓ 池田 香凜(九州大学 マス・フォア・インダストリ研究所)

▶ 学生モデレーター

- ✓ 藤江 勇仁(九州大学大学院マス・フォア・イノベーション連係学府)
- ✓ 寺田 怜央(九州大学大学院マス・フォア・イノベーション連係学府)

課題の説明

Weighted-signed-network におけるランキングスコア設計

➤ 与えられているもの: 重みつき有向グラフ

- ✓ Bitcoin取引者たちの相互評価から得られる。重みはGood～Badで10～-10の値を取る(生データは)。

➤ 作るもの: 取引者たちの高評価ランキングを与える関数

- ✓ 各人の評価基準の違いや評価数のばらつきを考慮した上で、どうやって客観的な指標にまとめるか、が問題の主幹
- ✓ 複数アカウントから恣意的な高評価を特定の取引者(もしくは相互)に与える、サクラの存在を考慮したい

評価アルゴリズムの構築

著者)

藤江 勇仁(九州大学大学院マス・フォア・イノベーション連係学府)

内海 忍(理化学研究所数理創造研究センター)

門脇 大我(九州大学大学院マス・フォア・イノベーション連係学府)

菅 陽佳(九州大学大学院マス・フォア・イノベーション連係学府)

西 皇龍(九州大学大学院数理学府)

※グラフ・図等は生成AI(Gemini, Claude, Copilot)を用いて生成

データセットから見る問題点の整理

➤ データセット

[SNAP: Signed network datasets: Bitcoin OTC web of trust network](#)

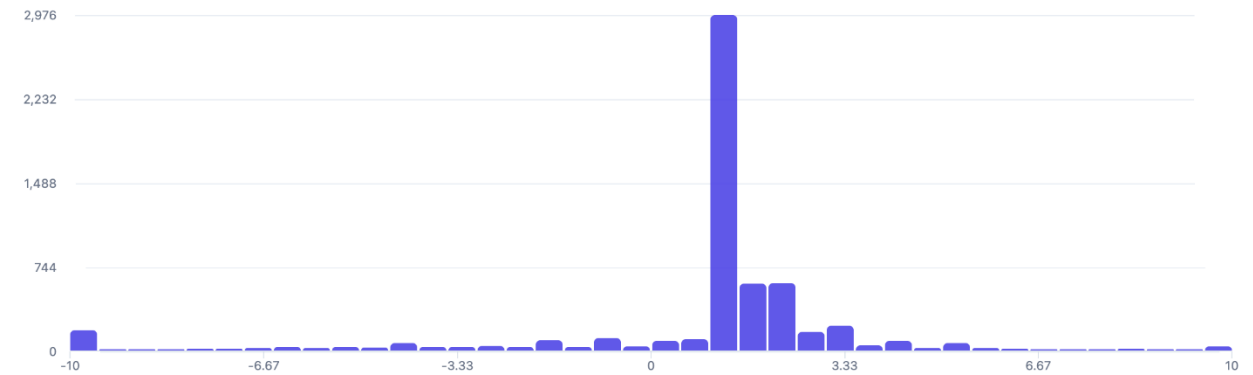
Bitcoinのプラットフォームで取引者間で行われた評価値をまとめたデータセット.

- 評価者 5881人
- 評価数 35,592件
- 評価幅 -10 ~ +10

やはり,

- 個々人の基準値のばらつき
- 受ける評価量のばらつき

等が観察できた.

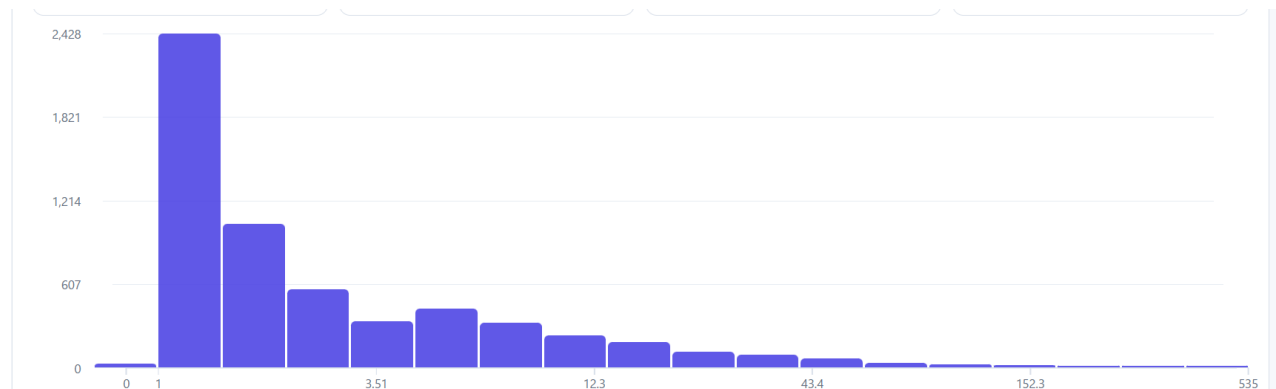


対象 5,858 ユーザー / 平均 0.73 / 範囲 -10 ~ 10 / 横軸 線形 / 全評価

図1: ユーザーが受け取る評価値の分布

観察① 受け取る好評/悪評の総量が異なる.

観察② 高評価の時は, +1 の評価をする人が多い.



対象 5,881 ユーザー / 平均 6.05 / 範囲 0 ~ 535 / 横軸 対数 / 全評価

図2: ユーザーが受け取る評価数の分布 (対数軸)

観察③ 受け取る評価が少ない(1件)のユーザーが大半を占める.

評価アルゴリズムの全景

➤ ノード u から v への評価点の次の3ステップで与える

1. 評価基準の差を補正する. 2. 評価件数の差を補正する. 3. 好評/悪評のバランスを取る.

例

好評の時に+1を付ける人が全体としては多いが、
+10だけを付けている人もいる。
さらに、評価幅も人によって違う。

解決策

好評/悪評それぞれを分離し、
個々の評価を全体の平均に寄せる。

例

★★★★★(2)と★★★★☆(400)
のようなレビューがあったときにどちら
が良いと評価するか？
→多くの人が後者を取ると予想。

解決策

評価件数と評価の平均値の両方を
”イイカンジ”に評価に落とし込む方法
を取りたい。→ベイズ収縮平均

例

良い評価の時は +1 が多く、
悪い評価の時は -10 が多い。
→評価 $+x$, $-x$ の強さは等しいのか？

解決策

好評の重みと悪評の重みが
同等になるように比率を整える。

(最初に好評/悪評を分離しているため、
最終的に足し合わせる時に調整すれば
よい。)

評価アルゴリズムの設計

[前提] まずは評価データの時系列は用いずに、最終的なネットワークを使用する

1. 個人の評価基準値を揃える → 好評/悪評を分けてそれぞれのスコア分布を標準化する.

理由：評価者には「甘め／辛口」の個人差がある。(平均のズレ)
また、極端な点数しかつけないというバイアス(分散のズレ)

$$\mu^+ = \mathbb{E}[R^+], \sigma^+ = \sqrt{\text{Var}(R^+)}$$

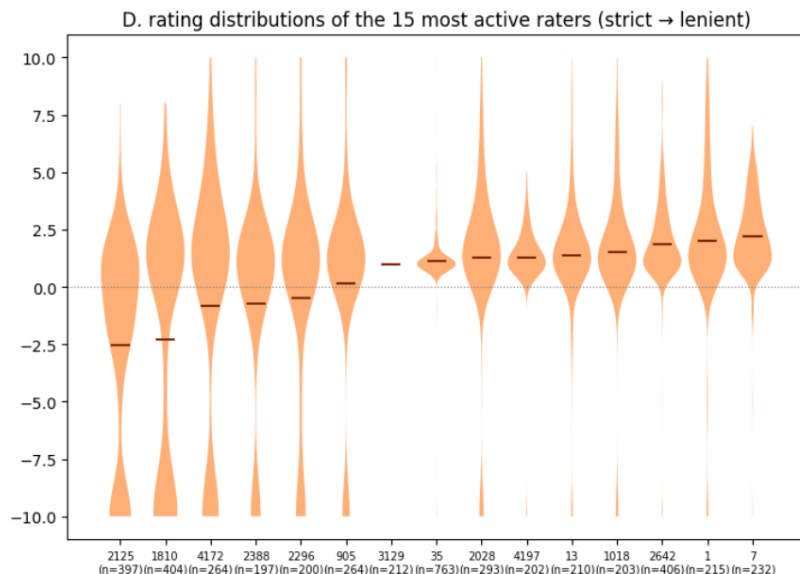
$$\mu^- = \mathbb{E}[R^-], \sigma^- = \sqrt{\text{Var}(R^-)}$$

R^+, R^- : 好評/悪評の評価値

μ^+, μ^- : 好評/悪評の評価値の平均値

σ^+, σ^- : 好評/悪評の評価値の標準偏差

例：評価している件数多い (Top15人) の評価のバラツキ



$$P_{uv}^{adj} = \begin{cases} \max \left(\min \left(\left(\frac{P_{uv} - m_u^+}{s_u^+} \right) \sigma^+ + \mu^+, 10 \right), 1 \right) & (s^+ \neq 0) \\ P_{uv} + (\mu^+ - m_u^+) & (s^+ = 0) \end{cases}$$

P_{uv} : ノード u から v への補正前の評価値

P_{uv}^{adj} : ノード u から v の補正後の評価値

m_u^+ : ノード u が他者に与えた評価値の平均

s_u^+ : ノード u が他者に与えた評価値の標準偏差 (好評の場合)

評価アルゴリズムの設計

2. 評価件数を加味したランク付け

➤ 手法: ベイズ収縮平均

各ノードが受けた評価の分布を $A_{uv} \sim N(\theta_v, \sigma^2)$,
評価全体の平均の分布を $\theta_v \sim N(\mu, \tau^2)$ としたときに
そのノード v が受けた, 件数を考慮した評価スコアを以下のように決定する.

$$S_v = \frac{\sum_{u \in I(v)} A_{uv} + C\mu}{n_v + C} = \frac{n_v}{n_v + C} \bar{x}_v + \frac{C}{n_v + C} \mu$$

ただし、 $\bar{x}_v = \frac{1}{n_v} \sum_{u \in I(v)} A_{uv}$ $C = \frac{\sigma^2}{\tau^2}$

➤ 解釈: 値が μ の仮想評価数を C 件だけ最初に加える

✓ 今回のデータでは $C \approx 3$ と推定される. これより評価数が少ないと評価が全体の平均に寄っていく.

評価アルゴリズムの設計

3. 好評・悪評のバランスの考慮

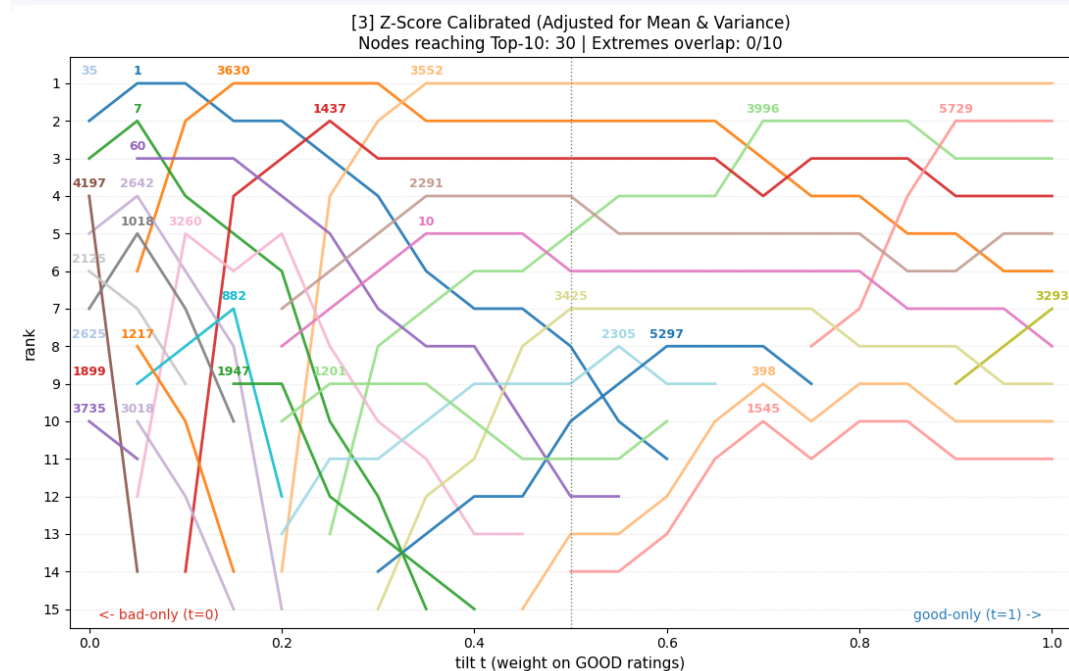
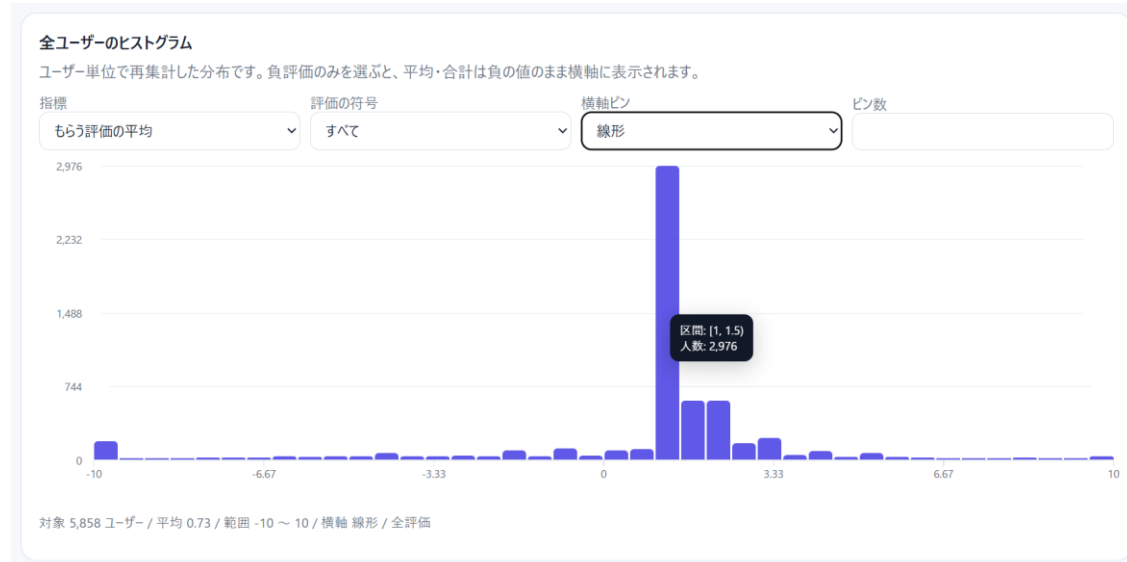
評価の重みを t とすることで、
スコア関数を次のように変更する。

$$S_v = \frac{tP(v) + (1-t)N(v) + C\mu}{n_v + C}$$

$$t = \frac{N}{P + N}$$

- ・ P は好評価の合計
- ・ N は悪評価の絶対値合計

この t を求めることで、悪評価数と好評価数との乖離を抑えることができる考えた。



実データへの評価アルゴリズムの適用(Bitcoin OTC ver)

➤ [SNAP: Signed network datasets: Bitcoin OTC web of trust network](#)について、上記に示した処理・ランク付けを行う。初めに、単純平均と単純合計でランキングを付けてみる。

単純平均

→少ない評価件数で、10点を得たノードが順位が高くなってしまふ

順位	ノード	スコア	件数	正件数	負件数
1	4823	10.0000	2	2	0
2	529	10.0000	1	1	0
3	814	10.0000	1	1	0
4	1122	10.0000	1	1	0
5	1261	10.0000	1	1	0
6	1326	10.0000	1	1	0
7	1340	10.0000	1	1	0
8	1501	10.0000	1	1	0
9	1545	10.0000	1	1	0
10	1663	10.0000	1	1	0

単純合計

→正の評価件数が圧倒的に多いノードが順位が高くなってしまふ。(談合、サクラによって順位を上げることが容易)

順位	ノード	スコア	件数	正件数	負件数
1	2642	1041.0000	412	411	1
2	35	1016.0000	535	535	0
3	1	801.0000	226	226	0
4	7	614.0000	216	216	0
5	4172	472.0000	222	211	11
6	1018	471.0000	179	179	0
7	2125	439.0000	180	180	0
8	4197	416.0000	203	203	0
9	4291	360.0000	158	156	2
10	13	341.0000	191	190	1

実データへの評価アルゴリズムの適用(Bitcoin OTC ver)

処理後のランキング

順位	ノード	スコア	件数	正件数	負件数
1	1	3.3853	226	226	0
2	1201	3.0826	58	57	1
3	3552	3.0487	16	14	2
4	7	2.7095	216	216	0
5	908	2.6775	54	54	0
6	60	2.6758	66	66	0
7	4	2.6002	54	54	0
8	2118	2.5554	25	25	0
9	3630	2.4991	23	23	0
10	1018	2.4840	179	179	0
11	3219	2.4821	47	47	0
12	1743	2.4524	11	11	0
13	2642	2.4492	412	411	1
14	1437	2.3966	9	9	0
15	1217	2.3905	81	81	0
16	23	2.3490	26	26	0
17	3018	2.3053	64	64	0
18	2125	2.3031	180	180	0
19	2305	2.2876	6	6	0
20	1690	2.2764	28	26	2

このアルゴリズムを適用すると、実際に順位が変遷していることが分かる。TOP20を見ると負の評価を1,2件もらっていても、ランクインしているノードが存在する。

単純平均におけるランキングで見られたような極端な少数の高評価を受けていた人や、単純合計で見られたような単に正の評価が多いだけの人の順位が落ちていることが観察できた。（特に、前者の人はかなり順位を落としていた。）

→ 少なくとも、単純平均/合計よりは妥当そうなランキングが得られていると予想。

別のデータセットでのランキング変動の観察

➤ [SNAP: Signed network datasets: Bitcoin Alpha web of trust network](#)についても同様に処理・ランク付けを行った。処理前のランキングは以下を参照。

単純平均

順位	ノード	スコア	件数	正件数	負件数
1	414	10.0000	2	2	0
2	418	10.0000	2	2	0
3	776	10.0000	1	1	0
4	782	10.0000	1	1	0
5	788	10.0000	1	1	0
6	789	10.0000	1	1	0
7	790	10.0000	1	1	0
8	791	10.0000	1	1	0
9	794	10.0000	1	1	0
10	795	10.0000	1	1	0

単純合計

順位	ノード	スコア	件数	正件数	負件数
1	1	758.0000	398	398	0
2	2	735.0000	205	205	0
3	3	610.0000	251	250	1
4	4	588.0000	201	201	0
5	5	390.0000	146	145	1
6	6	388.0000	139	139	0
7	7	370.0000	195	186	9
8	8	345.0000	134	134	0
9	9	321.0000	133	126	7
10	10	288.0000	164	163	1

実データへの評価アルゴリズムの適用(Bitcoin Alpha ver)

処理後のランキング

順位	ノード	スコア	件数	正件数	負件数
1	2	3.3994	205	205	0
2	20	3.3171	53	53	0
3	4	2.7708	201	201	0
4	23	2.6767	65	65	0
5	28	2.6528	54	54	0
6	71	2.6495	22	22	0
7	37	2.5868	46	46	0
8	6	2.5830	139	139	0
9	94	2.5673	18	18	0
10	18	2.4099	78	78	0
11	56	2.3975	33	33	0
12	74	2.3854	24	24	0
13	8	2.3758	134	134	0
14	160	2.3252	9	9	0
15	53	2.2961	38	38	0
16	3	2.2675	251	250	1
17	39	2.2114	53	53	0
18	213	2.2076	6	6	0
19	5	2.2002	146	145	1
20	59	2.1422	35	35	0

OTCの場合と同様に、ランキング変動が見られた。
やはり、極端な少数の高評価を受けていた人は上位には
入りにくいことが観察できる。

まとめ

- 以下の観点から評価アルゴリズムを構築した
 - ✓ 個人の評価のばらつきを揃える
 - 過剰な高い点数のみをつける人(サクラ？初心者？)をコミュニティ内の平均に揃えることができる
 - ✓ 評価件数を加味したランク付け
 - ✓ 好評・悪評のバランスの考慮
- 単純な指標によるランキングよりも妥当と考えられるランキングが得られた

スコアリング手法の公理化について

著者)

岩倉 康樹(九州大学大学院マス・フォア・イノベーション関係学府)

薄井 利基(九州大学大学院マス・フォア・イノベーション関係学府)

杉林 直樹(九州大学大学院数理学府)

目次

- ① 先行研究
- ② 公理化
- ③ 今後の課題

目標と到達方法

- 目標：所望の設計方針に基づいた客観的なスコアリング手法の確立。
- 到達方法：先行研究をもとにスコアリング方法を公理化し、公理を変えることで目的に応じたスコアの定義を行う。

本研究では、先行研究の調査やアイデアのブラッシュアップに生成AI（ChatGPT、Gemini）を使用した。

先行研究の手法の公理化

$M_+(X)$: $X := [0, 1] \times [-1, 1]$ の有限多重部分集合からなる集合。

主結果

関数 $\Gamma_0 : M_+(X) \rightarrow [-1, 1]$ で

$$\textcircled{1} \quad \Gamma_0(A + B) = \frac{|A| \Gamma_0(A) + |B| \Gamma_0(B)}{|A| + |B|},$$

$$\textcircled{2} \quad \Gamma_0([(f, w)]) = fw$$

を満たすものはただ一つである。

数式の意味：

- $\textcircled{1}$ 全ての辺を平等に考慮する。
- $\textcircled{2}$ 一つの辺しかない場合は信頼度と重みの積をスコアとする。

先行研究の紹介

*Kumar, Spezzano, Subrahmanian, and Faloutsos,
“Edge Weight Prediction in Weighted Signed Networks” (2016)*

先行研究のスコアの設計方針

- ① スコアリングは“goodness $g : V \rightarrow [-1, 1]$ ” を用いて行う。
- ② “fairness $f : V \rightarrow [0, 1]$ ” という尺度を導入する。
- ③ 全体スコアと近い評価を与えている人ほど fairness が高い。
- ④ fairness が高いノードほど全体スコアに影響を与えている (fairness で重み付け)。

$$g(v) := \frac{1}{\left| \{u : u \xrightarrow{e} v\} \right|} \sum_{u \xrightarrow{e} v} f(u)W(e)$$

Future Works

- ① 辺ごとの信頼度を考慮した後続研究もある：

*Kumar, Hooi, Makhija, Kumar, Faloutsos, and Subrahmanian,
“REV2: Fraudulent User Prediction in Rating Platforms” (2018)*

これに対して、同様の公理化・特徴づけができるのか？

- ② 先行研究の “fairness” の与え方は妥当なのか？

「全体スコアと近い評価を与えているほど fairness が高い」という方針は妥当なのか？

- ③ 他グループが行っていたことと組み合わせるとどうなるのか？

時系列データに対する新しいfairnessとgoodnessの提案

著者)

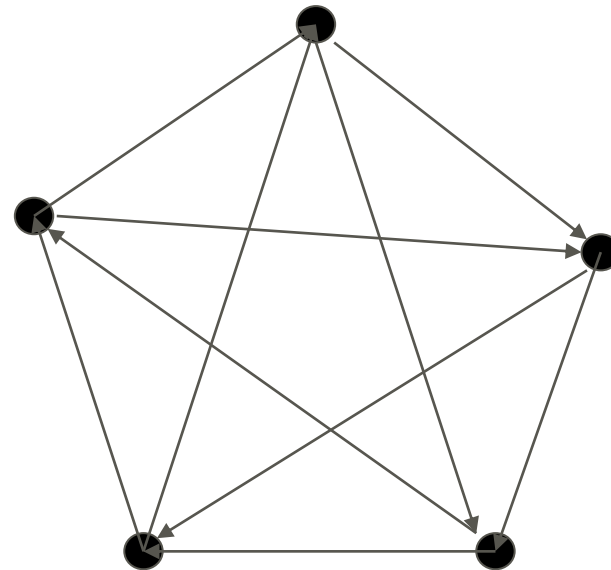
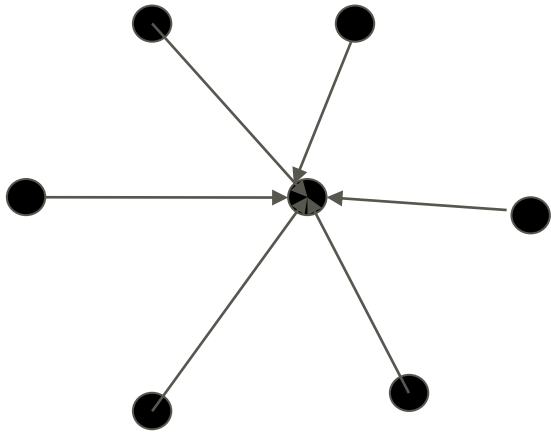
寺田 怜央(九州大学大学院マス・フォア・イノベーション関係学府)

時系列データ

集合 X , 部分集合 $V \subseteq X \times \mathbb{N}$, $E \subseteq X \times X \times \mathbb{N}$, 写像 $W: E \rightarrow [-1, 1]$ の組 (X, V, E, W) であって, 次の条件を満たすものについて考える.

- $u \in X, t, t', t'' \in \mathbb{N}$ が $(u, t), (u, t') \in V, t \leq t'' \leq t'$ を満たすならば $(u, t'') \in V$ である.
- $u, v \in X, t \in \mathbb{N}$ が $(u, v, t) \in E$ を満たすならば $(u, t), (v, t) \in V$ である.

- 時刻 t ごとに fairness と goodness を定義する.
- つぎのようなサクラの影響を弱める.



新しい fairness と goodness の提案

fairness $f: V \rightarrow [0, 1]$ と goodness $g: V \rightarrow [-1, 1]$ を次のように定義する.

$$f(u, t) := 1 - \frac{1}{N(t-1)} \left(\sum_{v \in \text{out}(u, t-1)} \frac{|w(u, v, t-1) - g(v, t-1)|}{2} + \sum_{v \in V(t) \setminus \text{out}(u, t-1)} 1 \right),$$

$$g(v, t) := \sum_{u \in \bigcup_{s=0}^{t-1} \text{in}(v, s)} \frac{\sum_{s=b(u)}^{t-1} w(u, v, s) f(u, s) \delta^{t-s-1}}{\sum_{s=b(u)}^{t-1} \delta^{t-s-1}}.$$

但し,

$$V(t) := \{x \in X \mid (x, t) \in V\},$$

$$N(t) := \#V(t),$$

$$\text{out}(u, t) := \{v \in X \mid (u, v, t) \in E\}$$

$$\text{in}(v, t) := \{u \in X \mid (u, v, t) \in E\}$$

$$b(u) := \min_{(t, u) \in V} t$$

$$w(u, v, t) := \frac{W(u, v, t)}{\sum_{x \in \text{out}(u, t)} |W(u, x, t)|}.$$

参考文献

- Srijan Kumar, Francesca Spezzano, VS Subrahmanian, and Christos Faloutsos. Edge weight prediction in weighted signed networks. In 2016 IEEE 16th international conference on data mining (ICDM), pages 221–230. IEEE, 2016.

時刻情報と評価件数情報を用いた スコアリング手法の修正

著者)

崎原 雄大(九州大学大学院マス・フォア・イノベーション関係学府)

藤井 秀一

力久 颯太(九州大学大学院数理学府)

既存研究 [Kumar et al, 2016]

符号重み付きネットワーク (Weighted Signed Network, WSN) $G = (V, E, W)$

V : ノード、 E : 有向辺、 $W(u \rightarrow v)$: u から v への辺の重み ($W(u \rightarrow v) \in [-1, 1]$)

Fairness

$$f(u) = 1 - \frac{1}{|\text{out}(u)|} \sum_{v \in \text{out}(u)} \frac{|W(u \rightarrow v) - g(v)|}{2} \in [0, 1]$$

Goodness

$$g(v) = \frac{1}{|\text{in}(v)|} \sum_{u \in \text{in}(v)} f(u) W(u \rightarrow v) \in [-1, 1]$$

既存研究の問題点

➤ 既存研究では時系列情報を利用していない

- ✓ 短期間に評価を連投する人はサクラの疑いがあるが、検出する仕組みがない
(評価した期間が長い人と、短い人を平等に扱っている)

➡ 評価期間が長いほど信頼できる評価者であると考えたい

➤ Fairness の算出に評価件数が考慮されていない

- ✓ 1件しか評価していない人と、数百件評価している人の評価の信頼度は同等か？

➡ 評価件数が多いほど公正な評価ができるようになる(＝評価が信頼できる)と考えたい

提案手法：活動実績の導入

1. 評価期間(時刻情報)による重みづけ

- 長い評価期間(他者を評価した期間)を持つ人ほど信頼できると仮定する

2. 評価件数による重みづけ

- 評価件数(した件数・された件数両方)が多い人ほど信頼できると仮定する

提案手法: 評価式

時刻符号重み付きネットワーク (Timed Weighted Signed Network, TWSN) $G = (V, E, W, T)$

V : ノード、 E : 有向辺、 $W(u \rightarrow v)$: u から v への辺の重み ($W(u \rightarrow v) \in [-1, 1]$)、

$T(u \rightarrow v)$: u から v に辺が張られた時刻 ($T(u \rightarrow v) \in [0, \infty)$)

修正Fairness

$$f(u) = \underbrace{\phi(\Delta T_{\text{out}}(u))}_{\text{評価期間}} \underbrace{\phi(|\text{out}(u)|)}_{\text{評価件数}} \left(1 - \frac{1}{|\text{out}(u)|} \sum_{v \in \text{out}(u)} \frac{|W(u \rightarrow v) - g(v)|}{2} \right) \in [0, 1]$$

修正Goodness

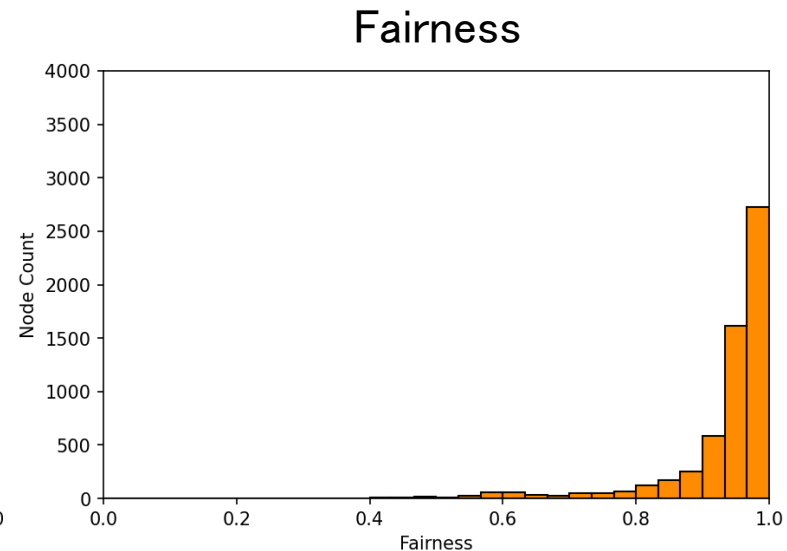
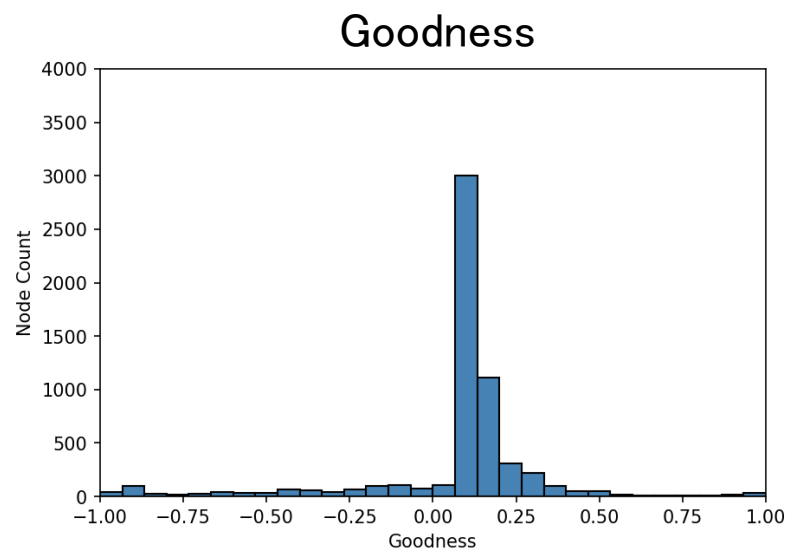
$$g(v) = \underbrace{\phi(\Delta T_{\text{in}}(v))}_{\text{評価期間}} \frac{1}{|\text{in}(v)|} \sum_{u \in \text{in}(v)} f(u) W(u \rightarrow v) \in [-1, 1]$$

- $\phi(x) = \frac{1}{1+e^{-x}}$ ($x \geq 0$) (重み付けのための関数)
- $\Delta T_{\text{out}}(u) = \max_{v \in \text{out}(u)} T(u, v) - \min_{v \in \text{out}(u)} T(u, v)$ (評価した期間)
- $\Delta T_{\text{in}}(v) = \max_{u \in \text{in}(v)} T(u, v) - \min_{u \in \text{in}(v)} T(u, v)$ (評価された期間)

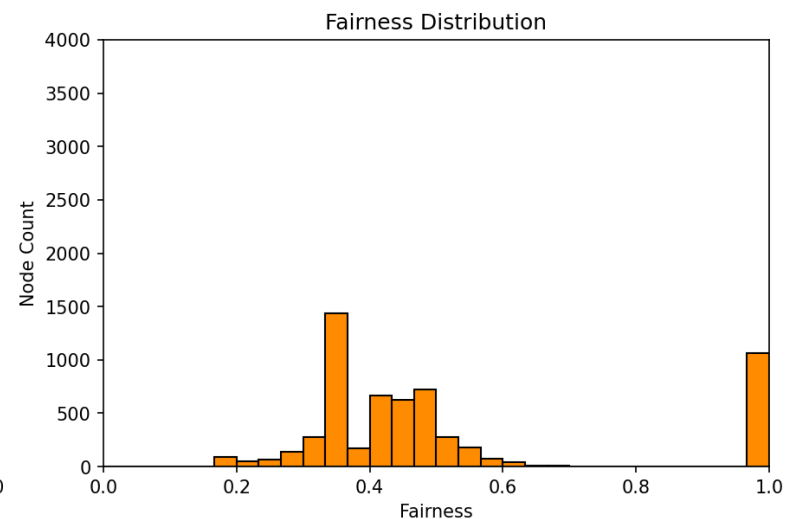
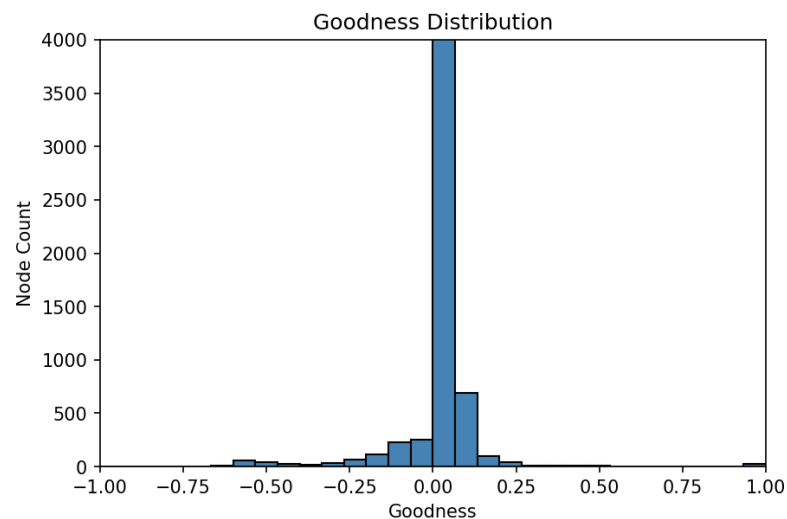
計算結果

➤ Bitcoin OTCデータセットにおけるGoodnessとFairnessの分布

修正前



修正後

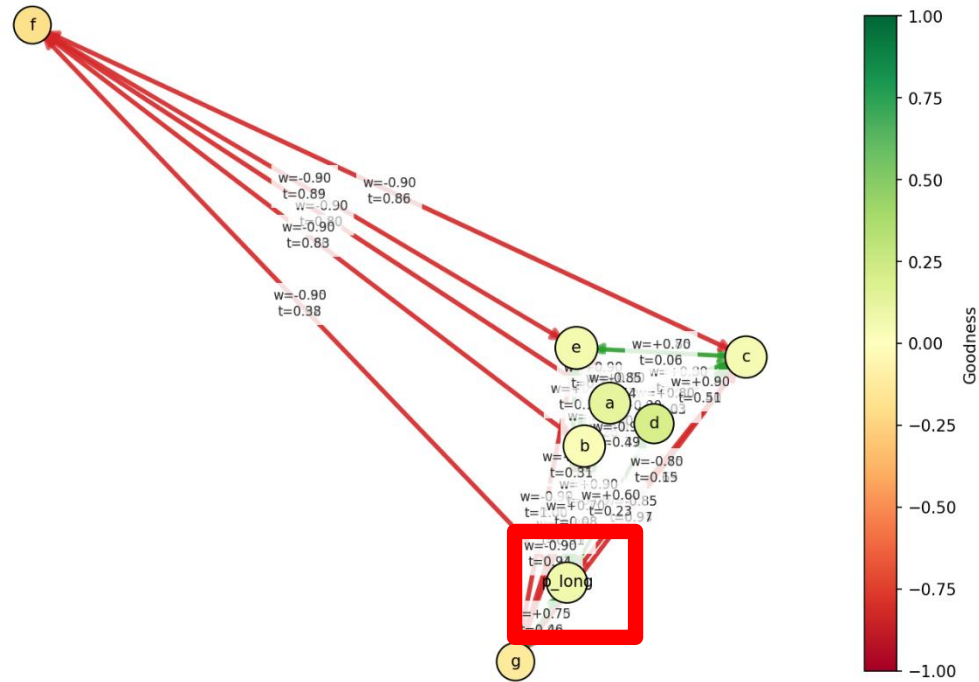


計算結果

➤ 単純な人工データを用いた計算結果

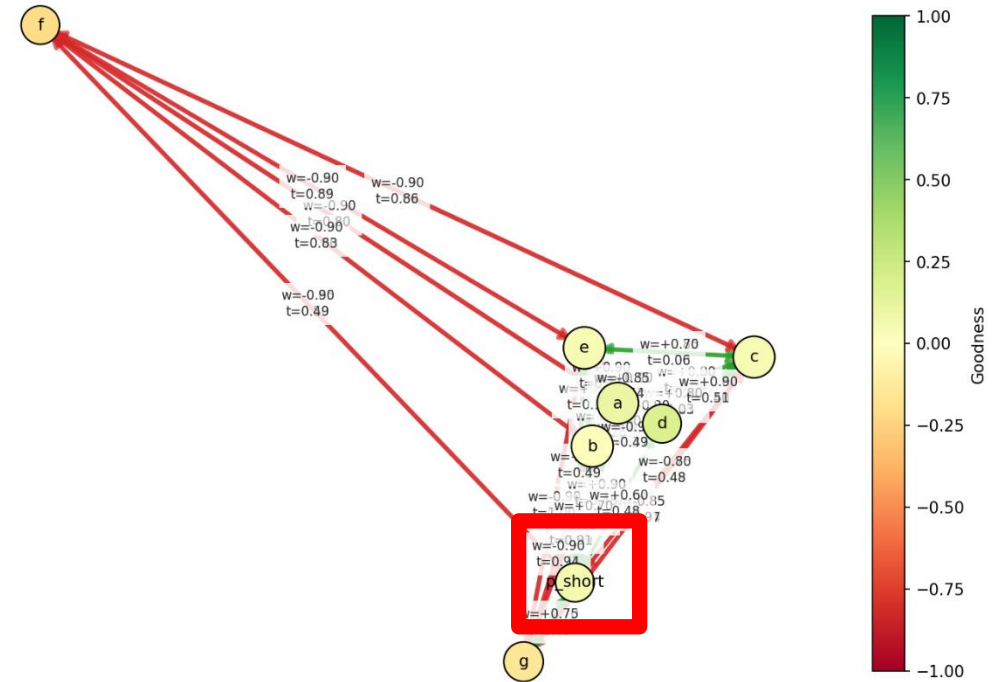
- ✓ 公正なユーザーの中に、不正者を一人追加したネットワークにおけるGoodness

Toy Example Graph (color=goodness, size=fairness, edge label=weight, time)



評価期間: 長

Toy Example Graph (color=goodness, size=fairness, edge label=weight, time)



評価期間: 短

今後の展望

➤ 様々なタイプの攻撃・悪人の検出

- ✓ “最初は善良に振る舞い、後で裏切る”タイプの悪人
- ✓ 他者の評価を模倣して多数の評価を行い信頼度を高め、その後特定の人に低評価を与える
- ✓ 特定の人々の評価を下げるために大量の低評価を与える
- ✓ → 評価頻度や評価パターンを考慮した重みづけ手法を検討する

参考文献

- Srijan Kumar, Francesca Spezzano, VS Subrahmanian, and Christos Faloutsos. Edge weight prediction in weighted signed networks. In 2016 IEEE 16th international conference on data mining (ICDM), pages 221–230. IEEE, 2016.