REV2: Fraudulent User Prediction in Rating Platforms

Srijan Kumar
Stanford University, USA
srijan@cs.stanford.edu

Bryan Hooi
Carnegie Mellon University, USA
bhooi@cs.cmu.edu

Disha Makhija
Flipkart, India
disha.makhija@flipkart.com

Mohit Kumar
Flipkart, India
k.mohit@flipkart.com

Christos Faloutsos
Carnegie Mellon University, USA
christos@cs.cmu.edu

V.S. Subrahmanian
Dartmouth College, USA
vs@dartmouth.edu

WSDM 2018

张峻伟
2019/5/9

本文的目的是为了识别欺诈性用户。由于缺乏培训标签，欺诈性和非欺诈性用户的比例不平衡，并缺乏培训标签，这项任务具有挑战性。

- 算法：我们提出三个指标，称为公平性，良好性和可靠性，分别对用户，产品和评级进行排名。 我们提出贝叶斯方法来解决冷启动问题并结合行为属性。 我们建议使用Rev2算法迭代计算这些指标。它适用于无人监督和监督设置。
- 理论保证：Rev2保证在有限的迭代次数中收敛，并具有线性时间复杂度。
- 有效性：Rev2在识别欺诈用户方面优于九种现有算法，通过五个评级网络上的五个实验进行，AUC≥0.85

现有的评级欺诈检测工作可分为基于网络的欺诈检测算法和基于行为的欺诈检测算法。

- 本文提出了量化用户、物品和评价的三个指标。我们将用户对项目的评分与时间戳建模为定向的二分图。从直觉上讲，公平的用户应该给出与产品，和好的产品应该得到高度肯定的可靠评级。显然，F(U)，G(P)，R(u，p)都是相互关联的，因此我们定义了这些内在度量应该满足的六个公理。关于彼此的关系。我们提出了三个满足公理的相互递推方程来计算这些度量的值。
- 解决了冷启动问题，增加了拉普拉斯平滑，以纳入默认的优先信念。
- 定义一个公理，建立了内在分数与行为特性之间的关系，并提出了一种贝叶斯技术，将用户、评级和产品的行为属性通过对异常行为的惩罚而纳入公式。

REV2 Formulation

二分评分图G =（U，R，P）是有向加权图

and $\mathtt{In(p)}$ be the set of ratings received by product p. So, $|\mathtt{Out(u)}|$ and $|\mathtt{In(p)}|$ represents their respective counts. The egonetwork of a user $u = \mathtt{Out(u)} \bigcup \{p | (u, p) \in Out(u)\}$ and that of product $p = \mathtt{In(p)} \bigcup \{u | (u, p) \in In(p)\}$.

Definition 1 [Identical ratings egonetworks]: Two users u_1 and u_2 are said to have identical ratings egonetworks if $|\mathtt{Out}(u_1)| = |\mathtt{Out}(u_2)|$ and there exists a one-to-one mapping $h : Out(u_1) \to Out(u_2)$ such that $\mathtt{score}(u_1, \mathtt{p}) = \mathtt{score}(u_2, \mathtt{h(p)})\ \forall (u_1, p) \in \mathtt{Out}(u_1)$. Similarly, identical ratings egonetworks of two products p_1 and p_2 can be defined in a similar manner using ratings the products receive.

Intrinsic Properties: Fairness, Goodness and Reliability

- Products vary in their quality, measured by a metric called ***goodness***. The goodness score is a single number indicating the most likely rating a fair user would give it. Intuitively, a good product would get several high positive ratings from fair users, and a bad product would receive high negative ratings from fair users. The goodness score $G(p)$ of a product p ranges from -1 (a very low quality product) to $+1$ (a very high quality product)
- Users vary in terms of their ***fairness*** that indicates how trustworthy it is. Fair users rate products without bias, i.e., they give high scores to high quality products, and low scores to bad products.
- Finally, ratings vary in terms of ***reliability***, which reflects how trustworthy it is. The reliability score $R(u, p)$ of a rating (u, p) ranges from 0 (untrustworthy) to 1 (trustworthy) $\forall (u, p) \in \mathcal{R}$.

Definition 2 [Identically reliable egonetworks]: We say that, two users u_1 and u_2 are said to have identically reliabile egonetworks if $|\texttt{Out}(u_1)| = |\texttt{Out}(u_2)|$ and there exists a one-to-one mapping $h : Out(u_1) \rightarrow Out(u_2)$ such that $\texttt{reliability}(u_1,\ \texttt{p}) = \texttt{reliability}(u_2,\ \texttt{h(p)})\ \forall (u_1, p) \in \texttt{Out}(u_1)$. Simiarly, identical ratings egonetworks of two products p_1 and p_2 can be defined in a similar manner by the ratings the products receive.

Five intuitive axioms

The relation between a product and the ratings that it receives.

AXIOM 1 (BETTER PRODUCTS GET HIGHER RATINGS). *If two products have identically reliable egonetworks and for one product, all the rating scores are higher, then quality of that product is more. Formally, for two products p_1 and p_2 have a one-to-one mapping $h : In(p_1) \to In(p_2)$ such that $R(u, p_1) = R(h(u), p_2)$ and* $\mathrm{score}(u, p_1) \geq \mathrm{score}(h(u), p_2)$ $\forall (u, p_1) \in In(p_1)$*, then* $G(p_1) \geq G(p_2)$.

AXIOM 2 (BETTER PRODUCTS GET MORE RELIABLE POSITIVE RATINGS). *If two products have identical rating egonetworks and for the first product, all positive ratings are more reliable and all negative ratings are less reliable than for the second product, then the first product has higher quality. Formally, if two products p_1 and p_2 have a one-to-one mapping $h : In(p_1) \to In(p_2)$ such that $R(u, p_1) \geq R(h(u), p_2) \forall (u, p_1) \in In^+(p_1)$ and $R(u, p_1) \leq R(h(u), p_2) \forall (u, p_1) \in In^-(p_1)$, then $G(p_1) \geq G(p_2)$.*

Proof: To prove axiom 1, let us take two products p_1 and p_2 that have a one-to-one mapping $h : In(p_1) \rightarrow In(p_2)$ such that $R(u, p_1) = R(h(u), p_2)$ and $\texttt{score(u,}\ p_1\texttt{)} \geq \texttt{score(h(u),}\ p_2\texttt{)}\ \forall (u, p_1) \in In(p_1)$.

From equation 2,

$$G(p_1) = \frac{\sum_{\texttt{(u,}p_1\texttt{)}\in\texttt{In(}p_1\texttt{)}} R(u, p_1) \cdot \texttt{score(u,}p_1\texttt{)}}{|\texttt{In(}p_1\texttt{)}|}$$

and

$$G(p_2) = \frac{\sum_{\texttt{(h(u),}p_2\texttt{)}\in\texttt{In(}p_2\texttt{)}} R(h(u), p_2) \cdot \texttt{score(h(u),}p_2\texttt{)}}{|\texttt{In(}p_2\texttt{)}|}$$

As the neighborhoods are identical with $|\texttt{In(}p_1\texttt{)}| = |\texttt{In(}p_2\texttt{))}|$ and $R(u, p_1) = R(h(u), p_2)$, so

$$G(p_1) - G(p_2) = \frac{\sum_{\texttt{(u,}p_1\texttt{)}\in\texttt{In(}p_1\texttt{)}} R(u, p_1) \cdot (\texttt{score(u,}p_1\texttt{)} - \texttt{score(h(u),}p_2\texttt{)})}{|\texttt{In(}p_1\texttt{)}|}$$

As $\texttt{score(u,}\ p_1\texttt{)} \geq \texttt{score(h(u),}\ p_2\texttt{)}$, so

$$G(p_1) - G(p_2) \geq \frac{\sum\limits_{(\mathtt{u},p_1)\in \mathtt{In}(p_1)} R(u, p_1)}{|\mathtt{In}(p_1)|}$$

As $R(u, p_1) > 0$ because all reliabilities are non-negative,

$$G(p_1) - G(p_2) \geq 0 \Rightarrow G(p_1) \geq G(p_2)$$

The other axioms have a very similar and straightforward proof.

The relation between a rating and the user and product it belongs to.

AXIOM 3 (RELIABLE RATINGS ARE CLOSER TO GOODNESS SCORES). *For two ratings by equally fair users, the rating with score closer to the product's gooness has higher reliability. Formally, if two ratings* (u_1, p_1) *and* (u_2, p_2) *are such that* $\mathrm{score}(u_1, p_1) = \mathrm{score}(u_2, p_2)$, $F(u_1) = F(u_2)$, *and* $|\mathrm{score}(u_1, p_1) - G(p_1)| \geq |\mathrm{score}(u_2, p_2) - G(p_2)|$, *then* $R(u_1, p_1) \leq R(u_2, p_2)$.

AXIOM 4 (RELIABLE RATINGS ARE GIVEN BY FAIRER USERS). *For two equal ratings to equal goodness products, the one given by more fair user has higher reliability. Formally, if two ratings* (u_1, p_1) *and* (u_2, p_2) *are such that* $\mathrm{score}(u_1, p_1) = \mathrm{score}(u_2, p_2)$, $F(u_1) \geq F(u_2)$, *and* $G(p_1) = G(p_2)$, *then* $R(u_1, p_1) \geq R(u_2, p_2)$.

the relation between a user and its ratings.

Axiom 5 (Fairer users give more reliable ratings). *For two users with equal number of ratings, if one has higher reliability for all its ratings than the other, then it has higher fairness. Formally, if two users u_1 and u_2 have a one-to-one mapping $h : Out(u_1) \rightarrow Out(p_2)$ such that $|Out(u_1)| = |Out(u_2)|$ and $R(u_1, p) \geq R(u_2, h(p))$ $\forall (u_1, p) \in Out(u_1)$, then $F(u_1) \geq F(u_2)$*

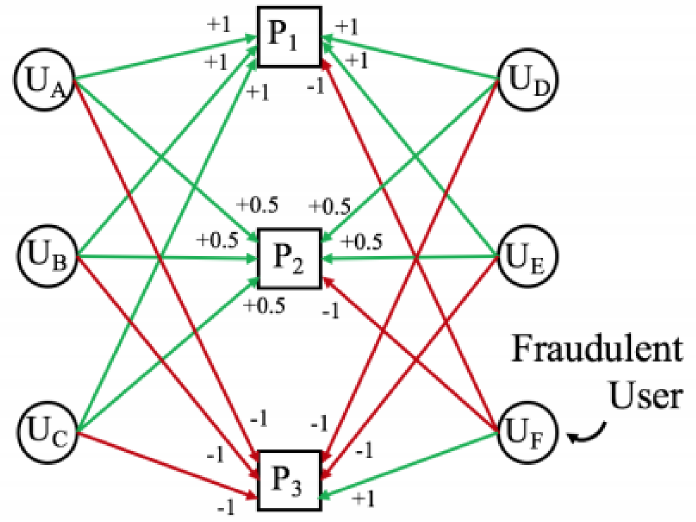


Figure 3: Toy example showing products (P_1, P_2, P_3), users (U_A, U_B, U_C, U_D, U_E and U_F), and rating scores provided by the users to the products. User U_F always disagrees with the ratings of others, so U_F is fraudulent.

Our proposed formulation

$$F(u) = \frac{\sum_{(u,p) \in \mathtt{Out(u)}} R(u,p)}{|\mathtt{Out(u)}|} \tag{1}$$

$$G(p) = \frac{\sum_{(u,p) \in \mathtt{In(p)}} R(u,p) \cdot \mathtt{score(u,p)}}{|\mathtt{In(p)}|} \tag{2}$$

$$R(u,p) = \frac{(\gamma_1 \cdot F(u) + \gamma_2 \cdot (1 - \frac{|\mathtt{score(u,p)} - G(p)|}{2}))}{\gamma_1 + \gamma_2} \tag{3}$$

Example:

$$R(U_F, P_1) = \frac{1}{2}\left(F(U_F) + (1 - \frac{|-1 - G(P_1)|}{2})\right)$$

Addressing Cold Start Problems

$$F(u) = \frac{\sum_{(u,p)\in \mathrm{Out(u)}} R(u,p) + \alpha_1 \cdot \mu_f}{|\mathrm{Out(u)}| + \alpha_1}$$

$$G(p) = \frac{\sum_{(u,p)\in \mathrm{In(p)}} R(u,p) \cdot \mathrm{score(u,p)} + \beta_1 \cdot \mu_g}{|\mathrm{In(p)}| + \beta_1}$$

Smoothing parameters α_1 and β_1 are non-negative integer psuedocounts, and μ_f and μ_g are prior beliefs for fairness and goodness scores of new nodes, respectively. The prior is the default score each user and product has if it didn't give or get any ratings.

Incorporating Behavioral

$$F(u) = \frac{\sum_{(u,p)\in \mathrm{Out(u)}} R(u,p) + \alpha_1 \cdot \mu_f + \alpha_2 \cdot \Pi_U(u)}{|\mathrm{Out(u)}| + \alpha_1 + \alpha_2}$$

$$R(u,p) = \frac{\gamma_1 \cdot F(u) + \gamma_2 \cdot \left(1 - \frac{|\mathrm{score(u,p)} - G(p)|}{2}\right) + \gamma_3 \cdot \Pi_R(u,p)}{\gamma_1 + \gamma_2 + \gamma_3}$$

$$G(p) = \frac{\sum_{(u,p)\in \mathrm{In(p)}} R(u,p) \cdot \mathrm{score(u,p)} + \beta_1 \cdot \mu_g + \beta_2 \cdot \Pi_P(p)}{|\mathrm{In(p)}| + \beta_1 + \beta_2}$$

Let, $\Pi_U(u)$ represent the behavior 'normality' score of a user u, with respect to some set of behavior properties. Lower (higher, resp.) score indicates more anomalous (normal, resp.) behavior.

The relation between behavior properties and other components

Axiom 6. *For two users with identically reliable egonetworks, the one with higher behavior score has higher fairness. Formally, if two users* u_1 *and* u_2 *have a one-to-one mapping* $h : Out(u_1) \rightarrow Out(p_2)$ *such that* $|Out(u_1)| = |Out(u_2)|$, $R(u_1, p) = R(u_2, h(p))\ \forall (u_1, p) \in Out(u_1)$, *and* $\Pi_U(u_1) \geq \Pi_U(u_2)$, *then* $F(u_1) \geq F(u_2)$.

. BIRDNEST [7]

calculates a Bayesian estimate of how much a user's properties deviates from the global population of all users' properties, and assigns a suspicion score $0 \leq BN_U(u) \leq 1$ to each user u. Then, normality score is simply $\Pi_U(u) = 1 - BN_U(u)$.

Theoretical Guarantees of Rev2

Let us first set up the basics that we need in order to prove the lemma and the two theorems. As a reminder, $F^t(u), G^t(p)$ and $R^t(u,p)$ represent the fairness score of user u, goodness score of product p and reliability score of rating r at the end of iteration t of Algorithm 1, for some input $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \gamma_3$. Recall that $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \gamma_3$ are non-negative integers. $F^\infty(u), G^\infty(p)$ and $R^\infty(u,p)$ are their final values after convergence for the same $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \gamma_3$.

At the end of iteration t of Algorithm 1, and by Equations 4, 5 and 6 in Figure 1, we get,

$$F^t(u) = \frac{\sum_{(u,p')\in \mathtt{Out(u)}} R^t(u,p') + \alpha_1 \cdot \mu_f + \alpha_2 \cdot \Pi_U(u)}{|\mathtt{Out(u)}| + \alpha_1 + \alpha_2} \quad (1)$$

$$R^t(u,p) = \frac{\gamma_1 \cdot F^{t-1}(u) + \gamma_2 \cdot (1 - \frac{|\mathtt{score(u,p)} - G^t(p)|}{2}) + \gamma_3 \cdot \Pi_R(u,p)}{\gamma_1 + \gamma_2 + \gamma_3} \quad (2)$$

$$G^t(p) = \frac{\sum_{(u',p)\in \mathtt{In(p)}} R^{t-1}(u',p) \cdot \mathtt{score(u',p)} + \beta_1 \cdot \mu_g + \beta_2 \cdot \Pi_P(p)}{|\mathtt{In(p)}| + \beta_1 + \beta_2} \quad (3)$$

$$F^\infty(u) = \frac{\sum\limits_{(u,p')\in \mathtt{Out(u)}} R^\infty(u,p') + \alpha_1 \cdot \mu_f + \alpha_2 \cdot \Pi_U(u)}{|\mathtt{Out(u)}| + \alpha_1 + \alpha_2} \tag{4}$$

$$R^\infty(u,p) = \frac{\gamma_1 \cdot F^\infty(u) + \gamma_2 \cdot (1 - \frac{|\mathtt{score(u,p)} - G^\infty(p)|}{2}) + \gamma_3 \cdot \Pi_R(u,p)}{\gamma_1 + \gamma_2 + \gamma_3} \tag{5}$$

$$G^\infty(p) = \frac{\sum\limits_{(u',p)\in \mathtt{In(p)}} R^\infty(u',p) \cdot \mathtt{score(u',p)} + \beta_1 \cdot \mu_g + \beta_2 \cdot \Pi_P(p)}{|\mathtt{In(p)}| + \beta_1 + \beta_2} \tag{6}$$

Lemma 1 (Lemma 1) *The difference between a product p's final goodness score and its score after the first iteration is at most 1, i.e.* $|G^\infty(p) - G^1(p)| \leq 1$. *Similarly,* $|R^\infty(u,p) - R^1(u,p)| \leq 3/4$ *and* $|F^\infty(u) - F^t(u)| \leq 3/4$.

Proof.

During initialization, all the ratings have reliability scores of 1, i.e. $R^0(u,p) = \Pi_R(u,p), \forall (u,p) \in \mathcal{R}$, and this is used to calculate the value of G^1.

Now, we will prove that $|G^\infty(p) - G^1(p)| \leq 1$.

From Equations 3 and 6, and substituting $t = 1$, we get,

$$|G^\infty(p) - G^1(p)| = |\frac{\sum\limits_{(u,p)\in \mathtt{In(p)}} R^\infty(u',p) \cdot \mathtt{score(u',p)} + \beta_1 \cdot \mu_g + \beta_2 \cdot \Pi_P(p)}{|\mathtt{In(p)}| + \beta_1 + \beta_2}$$

$$- \frac{\sum\limits_{(u,p)\in \mathtt{In(p)}} R^0(u',p) \cdot \mathtt{score(u',p)} + \beta_1 \cdot \mu_g + \beta_2 \cdot \Pi_P(p)}{|\mathtt{In(p)}| + \beta_1 + \beta_2}|$$

$$\Rightarrow |G^\infty(p) - G^1(p)| = |\frac{\sum\limits_{(u',p)\in \mathtt{In(p)}} (R^\infty(u',p) - R^0(u',p)) \cdot \mathtt{score(u',p)}}{|\mathtt{In(p)}| + \beta_1 + \beta_2}|$$

Since $|x + y| \leq |x| + |y|$, we have,

$$|G^{\infty}(p) - G^{1}(p)| \leq \frac{\sum\limits_{(u',p)\in\mathtt{In(p)}} |(R^{\infty}(u',p) - R^{0}(u',p)) \cdot \mathtt{score(u',p)}|}{|\mathtt{In(p)}| + \beta_1 + \beta_2}$$

Since $|x \cdot y| = |x| \cdot |y|$, we get,

$$|G^{\infty}(p) - G^{1}(p)| \leq \frac{\sum\limits_{(u',p)\in\mathtt{In(p)}} |(R^{\infty}(u',p) - R^{0}(u',p))| \cdot |\mathtt{score(u',p)}|}{|\mathtt{In(p)}| + \beta_1 + \beta_2}$$

Now $|(R^{\infty}(u',p) - R^{0}(u',p))| \leq 1$, as all reliability scores lie in [0,1], and $|\mathtt{score(u',p)}| < 1$. So we get,

$$|G^{\infty}(p) - G^{1}(p)| \leq \frac{\sum\limits_{(u',p)\in\mathtt{In(p)}} 1}{|\mathtt{In(p)}| + \beta_1 + \beta_2} = \frac{|\mathtt{In(p)}|}{|\mathtt{In(p)}| + \beta_1 + \beta_2} \leq 1 \quad \Rightarrow |G^{\infty}(p) - G^{1}(p)| \leq 1$$

Theorem 1 (Convergence Theorem) *The difference during iterations is bounded as* $|R^{\infty}(u,p) - R^{t}(u,p)| \leq (\frac{3}{4})^{t}, \forall (u,p) \in \mathcal{R}$. *As* t *increases, the difference decreases and* $R^{t}(u,p)$ *converges to* $R^{\infty}(u,p)$. *Similarly,* $|F^{\infty}(u) - F^{t}(u)| \leq (\frac{3}{4})^{t}, \forall u \in \mathcal{U}$ *and* $|G^{\infty}(p) - G^{t}(p)| \leq (\frac{3}{4})^{(t-1)}, \forall p \in \mathcal{P}$.

Proof. We will prove the convergence of the rating reliability values using mathematical induction.
From Equations 2 and 5,

$$|R^{\infty}(u,p) - R^{t}(u,p)| = |\frac{\gamma_1 \cdot F^{\infty}(u) + \gamma_2 \cdot (1 - \frac{|\mathtt{score(u,p)} - G^{\infty}(p)|}{2}) + \gamma_3 \cdot \Pi_R(u,p)}{\gamma_1+\gamma_2+\gamma_3}$$

$$- \frac{\gamma_1 \cdot F^{t-1}(u) + \gamma_2 \cdot (1 - \frac{|\mathtt{score(u,p)} - G^{t}(p)|}{2}) + \gamma_3 \cdot \Pi_R(u,p)}{\gamma_1+\gamma_2+\gamma_3}|$$

$$\Rightarrow |R^{\infty}(u,p) - R^{t}(u,p)| = \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot |\gamma_1 \cdot (F^{\infty}(u) - F^{t-1}(u)) + \gamma_2 \cdot (\frac{|\mathtt{score(u,p)} - G^{t}(p)| - |\mathtt{score(u,p)} - G^{\infty}(p)|}{2})|$$

As $|x| - |y| \leq |x - y|$, applying it to the goodness terms,

$$|R^{\infty}(u,p) - R^{t}(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot |\gamma_1(F^{\infty}(u) - F^{t-1}(u)) + \gamma_2 \frac{|G^{\infty}(p) - G^{t}(p)|}{2}|$$

Now, to simplify, we separately calculate $F^{\infty}(u) - F^{t-1}(u)$, using Equations 1 and 4:

$$F^{\infty}(u) - F^{t-1}(u) = \frac{\sum_{(u,p') \in \mathtt{Out(u)}} R^{\infty}(u,p') + \alpha_1 \cdot \mu_f + \alpha_2 \cdot \Pi_U(u)}{|\mathtt{Out(u)}| + \alpha_1 + \alpha_2} - \frac{\sum_{(u,p') \in \mathtt{Out(u)}} R^{t-1}(u,p') + \alpha_1 \cdot \mu_f + \alpha_2 \cdot \Pi_U(u)}{|\mathtt{Out(u)}| + \alpha_1 + \alpha_2}$$

$$\Rightarrow F^{\infty}(u) - F^{t-1}(u) = \frac{\sum_{(u,p') \in \mathtt{Out(u)}} (R^{\infty}(u,p') - R^{t-1}(u,p'))}{|\mathtt{Out(u)}| + \alpha_1 + \alpha_2}$$

Replacing this in the previous equation:

$$|R^\infty(u,p) - R^t(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (|\gamma_1 \frac{\sum\limits_{(u,p')\in \mathtt{Out(u)}} (R^\infty(u,p')-R^{t-1}(u,p'))}{|\mathtt{Out(u)}|+\alpha_1+\alpha_2} + \gamma_2 \frac{|G^\infty(p)-G^t(p)|}{2}|)$$

As $|a+b| \leq |a|+|b|$,

$$\Rightarrow |R^\infty(u,p) - R^t(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (|\gamma_1 \frac{\sum\limits_{(u,p')\in \mathtt{Out(u)}} (R^\infty(u,p')-R^{t-1}(u,p'))}{|\mathtt{Out(u)}|+\alpha_1+\alpha_2}| + |\gamma_2 \frac{|G^\infty(p)-G^t(p)|}{2}|)$$

As $|x \cdot y| = |x| \cdot |y|$ and again, as $|x+y| \leq |x|+|y|$,

$$\Rightarrow |R^\infty(u,p) - R^t(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (\gamma_1 \frac{\sum\limits_{(u,p')\in \mathtt{Out(u)}} |R^\infty(u,p')-R^{t-1}(u,p')|}{|\mathtt{Out(u)}|+\alpha_1+\alpha_2}| + |\gamma_2 \frac{|G^\infty(p)-G^t(p)|}{2}|) \tag{7}$$

Base case of induction.
We need to prove in the base case that, $|R^{\infty}(u,p) - R^{1}(u,p)| < 3/4$.
Substituting t = 1 in Equation 7,

$$|R^{\infty}(u,p) - R^{1}(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (\gamma_1 \frac{\sum_{(u,p')\in \texttt{Out(u)}} |R^{\infty}(u,p') - R^{0}(u,p')|}{|\alpha_1+\alpha_2+|\texttt{Out(u)}||} + \gamma_2 \frac{|G^{\infty}(p) - G^{1}(p)|}{2})$$

Trivially, $|R^{\infty}(r') - R^{0}(r')| < 1$, since reliability scores always lie in [0,1]. And from Lemma 1, $|G^{\infty}(p) - G^{1}(p)| < 1$. So,

$$|R^{\infty}(u,p) - R^{1}(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (\gamma_1 \frac{\sum_{r' \in \texttt{Out(u)}} 1}{|\alpha_1+\alpha_2+|\texttt{Out(u)}||} + \gamma_2 \cdot \frac{1}{2})$$

$$\Rightarrow |R^{\infty}(u,p) - R^{1}(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (\gamma_1 \frac{|\texttt{Out(u)}|}{|\alpha_1+\alpha_2+|\texttt{Out(u)}||} + \gamma_2 \cdot \frac{1}{2})$$

$$\Rightarrow |R^{\infty}(u,p) - R^{1}(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (\gamma_1 + \gamma_2 \cdot \frac{1}{2})$$

$$\Rightarrow |R^{\infty}(u,p) - R^{1}(u,p)| \leq \frac{\gamma_1}{\gamma_1+\gamma_2+\gamma_3} + \frac{\gamma_2}{\gamma_1+\gamma_2+\gamma_3} \cdot \frac{1}{2}$$

$$\Rightarrow |R^{\infty}(u,p) - R^{1}(u,p)| \leq 1 + \frac{1}{2}$$

$$\Rightarrow |R^{\infty}(u,p) - R^{1}(u,p)| \leq \frac{3}{4}$$

Induction Step.

We assume in the induction step that $|R^{\infty}(u,p) - R^{(t-1)}(u,p)| \leq \frac{3}{4}^{(t-1)}, \forall (u,p) \in \mathcal{R}$, which is consistent with the base case already.

We have to prove, $|R^{\infty}(u,p) - R^{t}(u,p)| \leq \frac{3}{4}^{(t)}$

We know from Equation 7, that

$$|R^{\infty}(u,p) - R^{t}(u,p)| \leq \frac{1}{\gamma_1 + \gamma_2 + \gamma_3} \cdot (\gamma_1 \frac{\sum\limits_{(u,p') \in \mathtt{Out(u)}} |R^{\infty}(u,p') - R^{t-1}(u,p')|}{|\mathtt{Out(u)}| + \alpha_1 + \alpha_2} | + | \gamma_2 \frac{|G^{\infty}(p) - G^{t}(p)|}{2} |)$$

We first find the bounds for $|G^{\infty}(p) - G^{t}(p)|$. From Equations 3 and 6, we get,

$$|G^{\infty}(p) - G^{t}(p)| = | \frac{\sum\limits_{(u',p) \in \mathtt{In(p)}} R^{\infty}(u',p) \cdot \mathtt{score(u',p)} + \beta_1 \cdot \mu_g + \beta_2 \cdot \Pi_P(p)}{|\mathtt{In(p)}| + \beta_1 + \beta_2} - \frac{\sum\limits_{(u',p) \in \mathtt{In(p)}} R^{t-1}(u',p) \cdot \mathtt{score(u',p)} + \beta_1 \cdot \mu_g + \beta_2 \cdot \Pi_P(p)}{|\mathtt{In(p)}| + \beta_1 + \beta_2} |$$

$$\Rightarrow |G^{\infty}(p) - G^{t}(p)| = | \frac{\sum_{(u',p) \in \mathtt{In(p)}} (R^{\infty}(u',p) - R^{t-1}(u',p)) \cdot \mathtt{score(u',p)}}{\beta_1 + \beta_2 + |\mathtt{In(p)}|} |$$

As $|x + y| \leq |x| + |y|$,

$$\Rightarrow |G^{\infty}(p) - G^{t}(p)| \leq \frac{\sum_{(u',p) \in \mathtt{In(p)}} |(R^{\infty}(u',p) - R^{t-1}(u',p)) \cdot \mathtt{score(u',p)}|}{|\beta_1 + \beta_2 + |\mathtt{In(p)}||}$$

As $|x.y| = |x|.|y|$,

$$\Rightarrow |G^{\infty}(p) - G^{t}(p)| \leq \frac{\sum_{(u',p)\in\texttt{In(p)}} |(R^{\infty}(u',p) - R^{t-1}(u',p))| \cdot |\texttt{score(u',p)}|}{|\beta_1 + \beta_2 + |\texttt{In(p)}||}$$

As $|\texttt{score(u',p)}| \leq 1, \forall (u',p) \in \mathcal{R}$ and $|(R^{\infty}(u',p) - R^{t-1}(u',p))| \leq \frac{3}{4}^{t-1}$, by induction assumption:

$$\Rightarrow |G^{\infty}(p) - G^{t}(p)| \leq \frac{\sum_{(u',p)\in\texttt{In(p)}} \frac{3}{4}^{t-1}}{|\beta_1 + \beta_2 + |\texttt{In(p)}||}$$

$$\Rightarrow |G^{\infty}(p) - G^{t}(p)| \leq \frac{3}{4}^{t-1} \cdot \frac{\sum_{(u',p)\in\texttt{In(p)}} 1}{|\beta_1 + \beta_2 + |\texttt{In(p)}||}$$

$$\Rightarrow |G^{\infty}(p) - G^{t}(p)| \leq \frac{3}{4}^{t-1} \cdot \frac{|\texttt{In(p)}|}{|\beta_1 + \beta_2 + |\texttt{In(p)}||}$$

$$\Rightarrow |G^{\infty}(p) - G^{t}(p)| \leq \frac{3}{4}^{t-1}$$

Substituting this back into Equation 7, and since by induction assumption $|R^{\infty}(u,p) - R^{t-1}(u,p)| \leq \frac{3}{4}^{t-1}$, we get

$$|R^{\infty}(u,p) - R^{t}(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (\gamma_1 \frac{\sum_{(u,p') \in \mathtt{Out(u)}} \frac{3}{4}^{t-1}}{|\mathtt{Out(u)}|+\alpha_1+\alpha_2} + \gamma_2 \frac{\frac{3}{4}^{t-1}}{2})$$

$$\Rightarrow |R^{\infty}(u,p) - R^{t}(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (\gamma_1 \cdot \frac{3}{4}^{t-1} \cdot \frac{\sum_{(u,p') \in \mathtt{Out(u)}} 1}{|\alpha_1+\alpha_2+|\mathtt{Out(u)}||} + \gamma_2 \cdot \frac{1}{2} \cdot \frac{3}{4}^{t-1})$$

$$\Rightarrow |R^{\infty}(u,p) - R^{t}(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (\gamma_1 \cdot \frac{3}{4}^{t-1} \cdot \frac{|\mathtt{Out(u)}|}{|\alpha_1+\alpha_2+|\mathtt{Out(u)}||} + \gamma_2 \cdot \frac{1}{2} \cdot \frac{3}{4}^{t-1})$$

$$\Rightarrow |R^{\infty}(u,p) - R^{t}(u,p)| \leq \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (\gamma_1 \frac{3}{4}^{t-1}) + \frac{1}{\gamma_1+\gamma_2+\gamma_3} \cdot (\gamma_2 \cdot \frac{1}{2} \cdot \frac{3}{4}^{t-1})$$

$$\Rightarrow |R^{\infty}(u,p) - R^{t}(u,p)| \leq \frac{\gamma_1}{\gamma_1+\gamma_2+\gamma_3} \frac{3}{4}^{t-1} + \frac{\gamma_2}{\gamma_1+\gamma_2+\gamma_3} \cdot \frac{1}{2} \cdot \frac{3}{4}^{t-1}$$

$$\Rightarrow |R^{\infty}(u,p) - R^{t}(u,p)| \leq \frac{3}{4}^{t-1} + \frac{1}{2} \cdot \frac{3}{4}^{t-1} = \frac{3}{4} \cdot \frac{3}{4}^{t-1} = \frac{3}{4}^{t}$$

$$\Rightarrow |R^{\infty}(u,p) - R^{t}(u,p)| \leq \frac{3}{4}^{t}$$

The proofs for $|F^{\infty}(u) - F^{t}(u)| \leq (\frac{3}{4})^{t}, \forall u \in \mathcal{C}$ follows from the above proof.

As t increases, $(\frac{3}{4})^{t} \to 0$ and $(\frac{3}{4})^{t-1} \to 0$, so after t iterations, $R^{t}(u,p) \to R^{\infty}(u,p), \forall (u,p) \in \mathcal{R}$. So after t iterations, the algorithm converges.

The entire of the above proof holds for any combination of parameters $(\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \gamma_3) \in \mathcal{C}$, therefore the entire algorithm converges for some t iterations. ■

Corollary 1 (Iterations till Convergence) *The number of iterations needed to reach convergence is at most* $2 + \lceil \frac{\log(\epsilon/2)}{\log(3/4)} \rceil$. *In other words, treating* ϵ *as constant, the number of iterations needed to reach convergence is bounded by a constant.*

Proof. Let $t = \lceil \frac{\log(\epsilon/2)}{\log(3/4)} \rceil$. By Theorem 1, after $t+1$ iterations, $\forall p \in \mathcal{P}$, $|G^{\infty}(p) - G^{t+1}(p)| \leq (\frac{3}{4})^t \leq (\frac{3}{4})^{\log_{3/4}(\epsilon/2)} = \epsilon/2$. Similarly $|G^{\infty}(p) - G^{t+2}(p)| \leq \epsilon/2 \cdot 3/4 \leq \epsilon/2$. Thus,

$$|G^{t+1}(p) - G^{t+2}(p)| = |G^{t+1}(p) - G^{\infty}(p) + G^{\infty}(p) - G^{t+2}(p)|$$

As $|x+y| \leq |x| + |y|$, we get

$$\Rightarrow |G^{t+1}(p) - G^{t+2}(p)| \leq |G^{t+1}(p) - G^{\infty}(p)| + |G^{\infty}(p) - G^{t+2}(p)|$$

$$\Rightarrow |G^{t+1}(p) - G^{t+2}(p)| \leq |G^{\infty}(p) - G^{t+1}(p)| + |G^{\infty}(p) - G^{t+2}(p)| \leq 2\epsilon/2 = \epsilon.$$

Similarly, for fairness, $|F^{t+1}(u) - F^{t+2}(u)| \leq \epsilon$ for all users u, and the same holds for reliability.

Thus, by the termination condition of Algorithm 1, convergence occurs after at most $t+2$ iterations. ∎

Experiment

- **Flipkart** is India's biggest online marketplace where users rate products. The ground truth labels are manually verified by review fraud investigators in Flipkart.
- **Bitcoin OTC** is a user-to-user trust network of Bitcoin users trading using OTC platform [14]. The network is made bipartite by splitting each user into a 'rater' with all its outgoing edges and 'product' with all incoming edges. The ground truth is defined as: benign users are the platform's founder and users he rated highly positively (≥ 0.5). Fraudulent users are the ones that these trusted users uniformly rate negatively (≤ -0.5).
- **Bitcoin Alpha** is the Bitcoin trust network of Alpha platform users [14]. Its ground truth is created similar to OTC, starting from the founder of this platform.
- **Epinions** network has two independent components—user-to-post rating network and user-to-user trust network [19]. Algorithms are run on the rating network and ground truth is defined using the separate trust network: a user is defined as benign if its total trust rating is $\geq +10$, and fraudulent if ≤ -10.
- **Amazon** is a user-to-product rating network [20]. Ground truth

Table 3: Unsupervised Predictions: The table shows the Average Precision values of all algorithms in unsupervised prediction of fraudulent and benign users across five datasets. The best algorithm in each column is colored blue and second best is light blue. Overall, REV2 performs the best or second best in 9 of the 10 cases. *nc* indicates 'no convergence'.

	Fraudulent user prediction					Benign user prediction				
	OTC	Alpha	Amazon	Epinions	Flipkart	OTC	Alpha	Amazon	Epinions	Flipkart
FraudEagle	93.67	86.08	47.21	*nc*	*nc*	86.94	71.99	96.88	*nc*	*nc*
BAD	79.75	63.29	55.92	58.31	79.96	77.41	68.31	97.19	97.09	38.07
SpEagle	74.40	68.42	12.16	*nc*	*nc*	80.91	82.23	93.42	*nc*	*nc*
BIRDNEST	61.89	53.46	19.09	37.08	85.71	46.11	77.18	93.32	98.53	62.47
Trustiness	74.11	49.40	40.05	*nc*	*nc*	84.09	78.19	97.33	*nc*	*nc*
REV2	96.30	75.29	64.89	81.56	99.65	92.85	84.85	100.0	99.81	42.83

Table 4: Supervised Predictions: AUC values of 10-fold cross validation to predict fradulent users with individual predictions as features in a Random Forest classifier. REV2 performs the best across all datasets. *nc* means 'no convergence'.

	OTC	Alpha	Amazon	Epinions	Flipkart
FraudEagle	0.89	0.76	0.81	*nc*	*nc*
BAD	0.79	0.68	0.80	0.81	0.64
SpEagle	0.69	0.57	0.63	*nc*	*nc*
BIRDNEST	0.71	0.73	0.56	0.84	0.80
Trustiness	0.82	0.75	0.72	*nc*	*nc*
SpEagle+	0.55	0.66	0.67	*nc*	*nc*
SpamBehavior	0.77	0.69	0.80	0.80	0.60
Spamicity	0.88	0.74	0.60	0.50	0.82
ICWSM'13	0.75	0.71	0.84	0.82	0.82
REV2	0.90	0.88	0.85	0.90	0.87

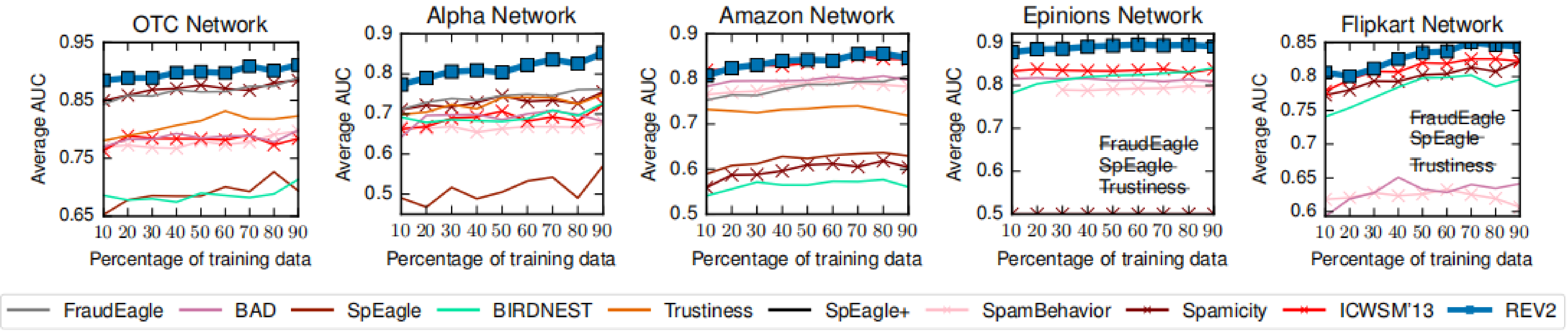
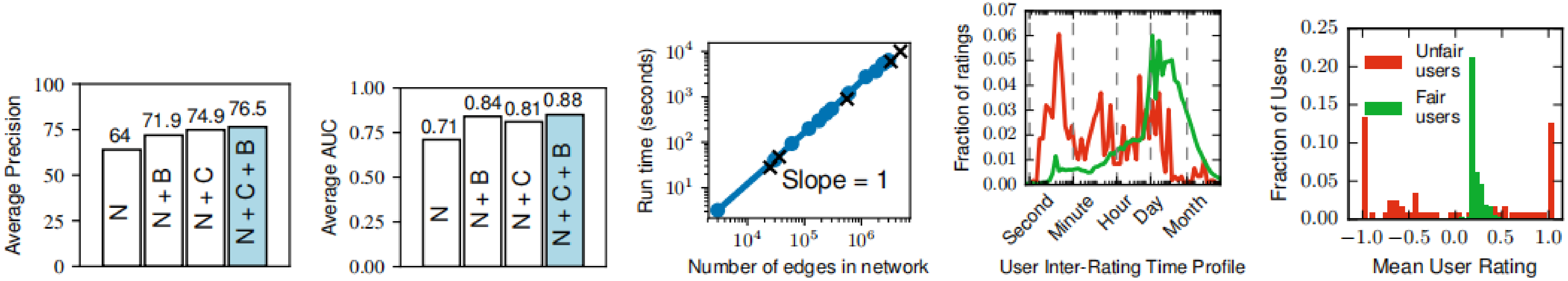


Figure 5: Variation of AUC for fraudulent user prediction with percentage of training data available for supervision. Rev2 consistently performs the best across all settings, and its performance is robust to the training percentage.



(a) Unsupervised Rev2 (b) Supervised Rev2 (c) Scalability of Rev2 (d) (e)

Figure 6: (a-b) Change in performance of Rev2 on Alpha network in (a) unsupervised and (b) supervised experiments when different components are used: network (N), cold start treatment (C) and behavioral (B). (c) Rev2 scales linearly—the running time increases linearly with the number of edges. (d-e) Fraudulent users identified by Rev2 in the OTC network are (d) faster in rating, and (e) give extreme ratings.