3-29-2018

A model for donation verification

Bin Fu
The University of Texas Rio Grande Valley

Fengjuan Zhu

John P. Abraham
The University of Texas Rio Grande Valley

A Model for Donation Verification

Bin Fu[1], Fengjuan Zhu[2], and John Abraham[1]

[1]Department of Computer Science
University of Texas Rio Grande Valley
bin.fu@utrgv.edu, john.abraham@utrgv.edu

[2]Department of Law
Shaoxing University, Shaoxin, P. C. China
zhufj@usx.edu.cn

October 21, 2018

Abstract

In this paper, we introduce a model for donation verification. A randomized algorithm is developed to check if the money claimed being received by the collector is $(1-\epsilon)$-approximation to the total amount money contributed by the donors. We also derive some negative results that show it is impossible to verify the donations under some circumstances.

Keywords: Donation, Verification Model, Approximation Algorithm, Randomization

1. Introduction

Worldwide billions of dollars are donated for charities. For example, United States alone gave over 335 billion dollars for philanthropy in 2013. When this much money is involved there would also be fraudsters who take advantage of ones generosity. Recognizing fraudulent practices, US Federal Trade Commission has given a number of things to check before giving to charity. The efficiency of a charitable organization is currently determined by the percentage of fund actually end up being used for intended purpose. CharityWatch [http://www.charitywatch.org/criteria.html] concludes 60% or greater spent on charitable programs and the remaining spent on overhead is acceptable. However, currently no algorithms are available to detects errors in reporting of monies donated. Donors merely trust the data provided by the charitable organizations or charts published by organizations such as Charity Navigator [http://www.charitynavigator.org]. Some research regarding charity donations and their management have been conducted in the academic community [1, 2, 3, 4, 5]. We have not seen any existing research about how donors check the amount of money received by the collector. It is essential to develop some algorithm that the donors and charitable organizations use to trust each other.

With the development of charity donations in the modern society, it becomes a more and more important social problem about charity donation system. In addition to establishing related laws, it is also essential to build up efficient auditing systems about charity donations, and apply big data technology to manage them. The progression in this direction will bring efficient and accurate methods for charity donations, which will improve our social reliability.

In this paper we develop a method that would allow us to verify monies received by charitable organizations. It would be difficult for every donor to verify each philanthropic organization. Our

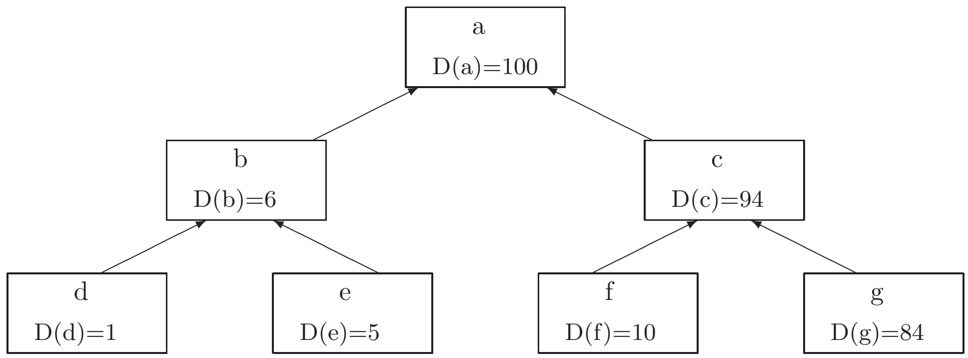

Figure 1: Donation Tree with D(.) Values

method is based on a randomized model thereby reducing the number of verifications. Using our algorithm, even if only a small percentage of the donors participate in the verification process, incorrect data given by the philanthropic organizations (cheating) can be detected. With just few steps a donor can verify if the money is used for intended purpose with a high degree of probability.

2. Models

Assume that there are n people who donated money. Person i donates s_i. In this model, we assume that each person checks his donation with probability $1-e^{-\lambda s}$ if he donated s amount money, where λ is fixed. This model means that a person will have larger chance to check his donation if he contributes more money. We define a donation tree.

Definition 1. We define a *donation tree*. For each leaf L in a donation tree, its donation value is defined to be $D(L)$. For node N in a tree T, define $D(N)$ to be the sum of values $D(.)$ in its leaves of the subtree with root at N. For each node D, function $V(N)$ is the amount money that the collector claims from the donors at the leaves of the subtree with root at N. An *error path* from a leaf to the root has a node N with $V(N) < V(N_1) + \cdots + V(N_k)$, where $N_1, \cdots, N_k$ are the children of N in T, and $V(N)$ be the vale saved in node N.

A donation tree without cheating should be the case $D(N) = V(N)$ for all nodes in the tree. We have the donation tree without cheating at Figure 1.

Let k be an integer at least 2. A k-donation tree is a donation tree such that each internal node has at most k children. The money donated from one donor is at a leaf. For every node saves N, the total money $D(N)$ of leave below it, and a value $V(N) \le D(N)$ to represents the amount of money the collector claiming to have received from the leaves. In the case $V(N) < D(N)$, it is considered a cheating from the collector.

3. Algorithm and Its Analysis

In this section, we develop an algorithm for this problem.

Lemma 2. *For each integer $k \ge 2$, A k-donation tree can be built in $O(n)$ time offline with depth $O(\log n)$. It also supports an $O(\log n)$ time for both insertion and deletion.*

Proof: A divide and conquer method can be used to build a donation tree of depth $O(\log n)$ with $O(n)$ time offline (the input of donations from n people are given). If it based on the structure of B+-tree, then it can support $O(\log n)$ time for both insertion and deletion. ▮

Protocol

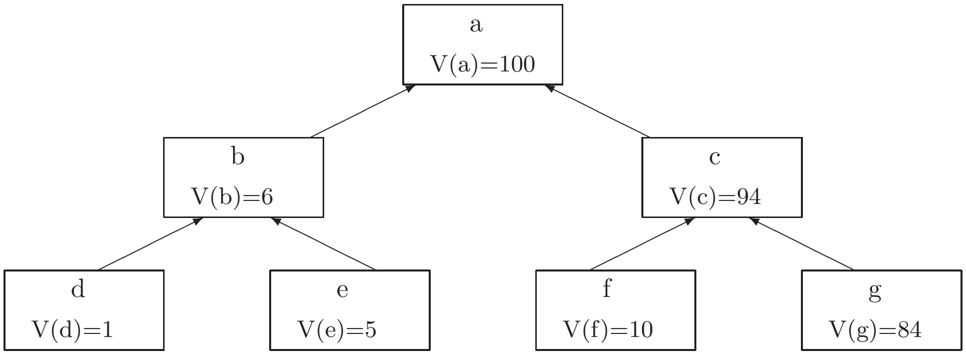

Figure 2: Donation Tree without Cheating

Collector:

Generate a file F that a donation tree.

Publish the file F.

donor m_i:

Check if his donation ($D(m_i) = V(m_i)$).

For each node N on the path from m_i to the tree root m_0, check if $V(N) = V(N_1) + \cdots + V(N_k)$, where $N_1, \cdots, N_k$ are all children of N.

Report error path if at least one of the two checks fails.

End of Protocol

Lemma 3. *Assume that the root N of the donation tree T has a value $V(N) < D(N)$. There are leaves $m_1, \cdots, m_k$ in the tree that all have error paths to root, and $D(m_1) + \cdots + D(m_k) \geq D(N) - V(N)$.*

Proof: We prove it by induction. It is trivial when the depth is 0. Assume that the statement is true for the depth at most d. Consider the depth $d+1$. Let N be a node of depth $d+1$ and has children $N_1, \cdots, N_k$.

Case 1. $V(N) < V(N_1) + \cdots + V(N_k)$, then every leaf has a error path.

Case 2. $V(N) \geq V(N_1) + \cdots + V(N_k)$. Let $N_{i_1}, \cdots, N_{i_t}$ be all of the nodes of $N_1, \cdots, N_k$ such that $V(N_{i_s}) < D(N_{i_s})$. We have

$$\begin{aligned} \sum_{j=1}^{t}(D(N_{i_j}) - V(N_{i_j})) &\geq \sum_{a=1}^{k}(D(N_a) - V(N_a)) && (1) \\ &\geq D(N) - V(N). && (2) \end{aligned}$$

We note that for each $a \in \{1, 2, \cdots, k\} - \{i_1, \cdots, i_t\}$, $D(N_a) - V(N_a) \leq 0$.

By induction hypothesis, for each i_j, there are leave $l_{i_j,1}, \cdots, l_{i_j,u}$ under the subtree with root at N_{i_j} such that $D(l_{i_j,1}) + \cdots + D(l_{i_j,u}) \geq D(N_{i_j}) - V(N_{i_j})$. Let $H_j = \{l_{i_j,1}, \cdots, l_{i_j,u}\}$ for $j = 1, 2, \cdots, t$. Let H be the set of all leave $m_i \in H_1 \cup \cdots H_t$, we have $\sum_{m_i \in H} D(m_i) \geq D(N) - V(N)$.

∎

3.1. Random Verification with Exponential Distribution

In this section, we consider the case that donor join the verification by following exponential distribution. The people who donate more money have higher probability to do the verification than the people who donate less money.

Theorem 4. *Assume integer $k \geq 2$ and there are n donors. Each verify takes $O(kh$ steps, and reports error if it is not an $(1-\epsilon)$-approximation in the report with probability at least $1-\delta$, where $\delta = O(e^{-\lambda\epsilon M})$ and h is the depth of the tree.*

Proof: Let $M = D(R)$ where R is the root of the donation tree. Assume that there is at least ϵM error ($D(R) - V(R) \geq \epsilon D(R)$).

Let $m_1, m_2, \cdots, m_t$ be nodes with error paths to root of the tree, and have $\sum_{i=1}^{t} m_i \geq \epsilon M$ by Lemma 3.

If one of $m_1, m_2, \cdots, m_t$ checks its path to the root, then an error (or cheating) can be detected. Therefore, this problem becomes to compute the probability that none of $m_1, m_2, \cdots, m_t$ does his verification.

The probability that none of them checks is at most $e^{-\lambda m_1} \cdot e^{-\lambda m_2} \cdots e^{-\lambda m_t} \leq e^{-\lambda\epsilon M}$. ∎

3.2. An Implementation with B-Tree

A donation tree can be implemented with a B-tree that supports $O(\log n)$ time for searching, insertion, and deletion. When an new leave is inserted, we can update all $V(N)$ for node N affected in $O(\log n)$ time. Similarly, When an new leave is deleted, we can update all $V(N)$ for node N affected in $O(\log n)$ time.

3.3. Uniform Random Verification

In this section, we consider the case that donor join the verification by following uniform distribution.

Theorem 5. *Assume that each person donates the money in the range $[1, a]$, each donor participates in the verification with probability at least δ. Then it takes $O(\log_k n)$ steps, and reports error with probability at least $1 - (1-\delta)^{\lceil \frac{\epsilon M}{a} \rceil}$.*

Proof: Assume that M is the total amount of money donated by all the people. If it is not an $(1-\epsilon)$-approximation, then there are at least $\frac{\epsilon M}{a}$ error paths corresponding to at least $k = \lceil \frac{\epsilon M}{a} \rceil$ donors. With probability at most $(1-\delta)^k = (1-\delta)^{\lceil \frac{\epsilon M}{a} \rceil}$, none of them will attend the verification. Therefore, with probability at least $1 - (1-\delta)^{\lceil \frac{\epsilon M}{a} \rceil}$, the error of the report will be detected. ∎

3.4. Multiple Verification Regions

In this section, we show the verification in several region. If each person donates amount in the range $[a_0, a]$. The interval is partitioned into $[a_0, a_1), [a_1, a_2), \cdots, [a_{k-1}, a_k]$. We assume people different region have different probability to participate the verification.

Theorem 6. *Assume that $[a_0, a_1), [a_1, a_2), \cdots, [a_{k-1}, a_k]$ form a partition for $[a_0, a]$ with $a_i + 1 \leq a_i(1+\delta)$ for $i = 0, 1, 2, \cdots, k-1$. Let $I_j = [a_j, a_{j+1})$ if $j < k$, and $I_k = [a_{k-1}, a_k]$. Let p_j be the probability that a person with donation range in I_j verifies. Then there is a verification protocol such that with probability at most $\sum_{j=0}^{k-1}(1-p_j)^{\epsilon M_j/(1+\delta)}$ to fail to check $1-\epsilon$ approximation, where M_j is the total amount of donation with each donation in I_j. Furthermore, the verification time is $O(\log n + k)$.*

Proof: Use one verification tree T_j for each I_j. Form a tree T by linking $T_1, \cdots, T_{k-1}$ as children. It follows from Theorem 5. ∎

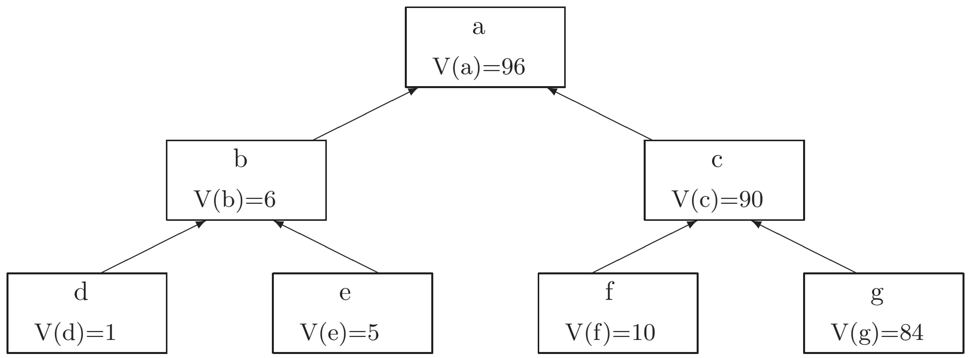

Figure 3: Donation Tree with Cheating. Both nodes f and g can find the cheating problem at their paths to the root. For example, at node c, $V(c) < V(f) + V(g)$.

4. Impossibilities of Verification

In this section, we show that it is impossible to use uniform probability to do donation verification. We also prove that it is impossible to do verification if negative items are allowed.

Theorem 7. *There is no randomized algorithm fail to detect the cheating from collector with probability at most δ if every donor checks his donation with probability at most δ.*

Proof: Let $k = 9$. Imagine the collector receives M amount money with $\frac{M}{k}$ from one donor A. He releases a document that includes all the money from the others except A. If A does not verify it, it should be all correct without any error. Therefore, with probability at most δ, the verification fails. ∎

Theorem 8. *There is no randomized algorithm fail to detect the cheating with negative donation allowed from collector with probability at most δ if one donor checks his donation with probability at most δ.*

Proof: Let the sum of $n-2$ donors $m_1, \cdots, m_{n-2}$ be equal to M. Let the donor $n-1$ contributes 1 or 0, and donor n contributes $-M$. Consider the first case that donor $n-1$ contributes 1. The total is equal to 1.

Consider the second case that donor $n-1$ contributes 0. The total is equal to 0. If that donor $n-1$ takes probability at most δ to do verification, then we have probability at most δ to make the difference of the two cases. ∎

5. Conclusions

In this paper, we develop a protocol for the donation verification under some probabilistic assumption. It only expects the donors follow certain probabilistic distribution to attend verification, and takes $O(\log n)$ steps for each donor.

References

[1] Adam Wociechowski, Models of Charity Donations and Project Funding in Social Networks, Lecture Notes in Computer Science 5872, pp. 454-463, 2009.

[2] Robert Cooter, and Brian J Broughman, Charity, Publicity, and the Donation Registry, The Economists' Voice. Volume 2, Issue 3, pp. 1553-3832, 2005.

[3] Sampath Kumar Ranganathan and Walter H. Henley, Determiniants of charitable donation intentions: a structural equation model, International Journal of Nonprofit and Voluntary Sector Marketing, DOI: 10.1002/nvsm.297, 2007.

[4] G. Dougla Olsen, John W. Pracejus, and Norman R. Brown, When profit equals price: consumer confusion about donation amounts in cause-related marketing. Journal of Public Policy & Marketing: Vol. 22, No. 2, pp. 170-180, 2003.

[5] Rogers Bennett, Factors influencing donation switching behaviour among charity supporters: an empirical investigation, Journal of Customer Behaviour, Volume 8, Number 4, pp. 329-345, 2009.