Stability of Weighted Majority Voting under Estimated Weights

Extended Abstract

Shaojie Bai
Zhejiang University
Hangzhou, China
white.shaojie@gmail.com

Dongxia Wang*
Zhejiang University
Hangzhou, China
wangdxee@gmail.com

Tim Muller
University of Nottingham
Nottingham, United Kingdom
tim.muller@nottingham.ac.uk

Peng Cheng*
Zhejiang University
Hangzhou, China
saodiseng@gmail.com

Jiming Chen
Zhejiang University
Hangzhou, China
cjm@zju.edu.cn

ABSTRACT

Weighted Majority Voting (WMV) is a well-known decision making rule. The weights of sources are determined by the probabilities that sources provide accurate information (*trustworthiness*). However, in reality, the trustworthiness is usually not a known quantity to the decision maker − they have to rely on an estimate called *trust*. An algorithm that computes trust is called *unbiased* when it has the property that it does not systematically overestimate or underestimate the trustworthiness. To formally analyze the uncertainty to the decision process brought by such unbiased trust values, we introduce and analyze two important properties of WMV: *stability of correctness* and *stability of optimality*. We also provide an overview of how sensitive decision accuracy is to the changes in trust and trustworthiness.

KEYWORDS

Weighted Majority Voting; Trust; Trustworthiness; Stability

ACM Reference Format:
Shaojie Bai, Dongxia Wang*, Tim Muller, Peng Cheng*, and Jiming Chen. 2023. Stability of Weighted Majority Voting under Estimated Weights: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Weighted Majority Voting has long been a popular decision-making mechanism, which has been applied in a variety of domains ranging from voting, crowdsourcing, classification to trust systems and even distributed systems [14, 15, 17, 19, 21]. In WMV, decisions are derived based on aggregating the feedback from a collection of sources. Each source is assigned a weight and that weight depends on how trustworthy the source is in providing the feedback that corresponds to the correct decision, denoted as *trustworthiness*, which is usually modeled as a probability value [12, 18]. A decision-maker usually has to resort to an estimation or a belief about trustworthiness, denoted as *trust* that may not equal *trustworthiness* [20]. We focus on the question: is WMV able to maintain a tolerant level of decision incorrectness with the inaccuracy in the estimation of trustworthiness bounded, meaning having certain levels of stability w.r.t the inaccurate estimation?

We propose a formal analysis of the stability properties of WMV. Firstly, we study how sensitive the decision accuracy is to the changes in source trust and trustworthiness, with both the arguments taking fixed values. Secondly, we consider unbiased estimation of trustworthiness[1] and its influence to WMV. We define *Stability of Correctness*, which measures whether the decision accuracy in belief equals that in reality, and prove WMV has absolute stability. Besides, we define *Stability of Optimality*, which measures the gap in decision accuracy between the cases where trust is used and where trustworthiness is used. We prove the degradation in decision accuracy caused by the incorrect but unbiased trust is relatively tightly bounded. That is, decision accuracy with unbiased trust will not be too far off the theoretically determined value.

2 WEIGHTED MAJORITY VOTING AND PARAMETER SENSITIVITY

2.1 Formal Framework

We outline a formal framework to support our study of the stability of WMV. Consider a decision-making scenario, a decision maker is faced with multiple possible decisions O=$\{o_1, o_2, \dots\ o_L\}$, and only one of them is correct denoted as C. The decision maker receives feedback $\boldsymbol{f}$:$\boldsymbol{f}$=$(f_1, \dots, f_n)$ from a set of sources: $\mathcal{S}$=$\{s_1, .., s_n\}$, where f_i denoting an outcome of random variable F_i with $\boldsymbol{F}$=$(F_1, \dots, F_n)$. For WMV, we assume a one-to-one correspondence between the feedback that suggests the correct decision and the correct decision itself, and denote F_i=C iff f_i suggests correctly, $F_i \in O$. For source i, let $\mathbf{Pr}(F_i$=$C)$=p_i be its trustworthiness and $\boldsymbol{p}$=$(p_1, \dots, p_n)$, and an estimation of it is denoted as $\hat{p}_i$ (trust), with $\hat{\boldsymbol{p}}$=$(\hat{p}_1, \dots \hat{p}_n)$.

Definition 1 (Weighted Majority Voting $\mathcal{D}_W$). *Given a set of n sources $\mathcal{S}$, their trustworthiness $\boldsymbol{p}$ and independent feedback $\boldsymbol{f}$, $\mathcal{D}_W$ makes decisions via the function [5, 16, 17]:*

$$\mathcal{D}_W(\boldsymbol{f}) = argmax_{o \in O} \left(\sum_{i:f_i=o} w_i \cdot f_i \right) \tag{1}$$

where $p_i \geq 0.5$, $f_i \in O$, and w_i=$\log(p_i/1-p_i)$.

* Corresponding author.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

[1]Generally, the estimation error always exists, but it is relatively small and can be zero on average with sufficient data in a statistical way [4, 9]

WMV has been proven to be optimal when trustworthiness $\boldsymbol{p}$ is used for decision making [17]. Trustworthiness is usually unknown in practice and the weight assigned to each source depends on trust instead: $w_i = \log(\hat{p}_i/1-\hat{p}_i)$. We introduce the probability of WMV's making correct decisions (decision accuracy) as [6, 15]:

$$\mathbf{Pr}(\mathcal{D}_W(f) = C) \triangleq \omega(\hat{\boldsymbol{p}}, \boldsymbol{p}) \tag{2}$$

In Equation 2, the first parameter of the function $\omega()$ represents the value used for decision-making, while the second represents the value for computing the probability of deciding correctly. Both the parameters can be either trust or trustworthiness, resulting in different meanings of decision accuracy. The quantity $\omega(\boldsymbol{p}, \boldsymbol{p})$ denotes the "ideal" decision accuracy, where the decision maker magically knows the trustworthiness values. The quantity $\omega(\hat{\boldsymbol{p}}, \boldsymbol{p})$ denotes the "practical" decision accuracy, where the decision maker decides with trust, but the accuracy he actually achieves depends on trustworthiness. The quantity $\omega(\hat{\boldsymbol{p}}, \hat{\boldsymbol{p}})$ denotes the decision accuracy "in belief", which represents what the decision maker thinks he can obtain while the actual accuracy may be different.

2.2 Parameter Analysis

In this section, we analyze how changes in the values of trust and trustworthiness influence the decision accuracy or the correctness of WMV. Firstly, we analyze the case where the parameter used for making decisions and that for computing accuracy are equal. There are two different rationales for doing this, i.e., the "ideal" decision accuracy $\omega(\boldsymbol{p}, \boldsymbol{p})$ and the decision accuracy "in belief" $\omega(\hat{\boldsymbol{p}}, \hat{\boldsymbol{p}})$, but the mathematics is identical for both.

Lemma 1. *Let $f(p_i)$=$\omega(\boldsymbol{p}, \boldsymbol{p})$, where p_j is constant for $j \neq i$. The function $f(p_i)$ is a piecewise linear non-decreasing convex function.*

Meanwhile, when multiple sources vary, the accuracy is also a piecewise non-decreasing function. Generally, if a source is more trustworthy, the decision accuracy increases with a positive second derivative.

With the help of Lemma 1, we can analyze the cases where trustworthiness and trust are not identical. When trustworthiness varies with trust value fixed, the line or surface of $\omega(\hat{\boldsymbol{p}}, \boldsymbol{p})$ corresponds to one of the segments from the piecewise function in Lemma 1.

When trust varies with trustworthiness value fixed, the accuracy $\omega(\hat{\boldsymbol{p}}, \boldsymbol{p})$ is a discontinuous staircase function consisting of flat plateaus, achieving maximum at the plateau containing $\hat{\boldsymbol{p}}$=$\boldsymbol{p}$. An insight is that the nearby points are more likely to be on the same plateau, meaning a small estimation deviation may be unlikely to affect the accuracy.

3 STABILITY ANALYSIS

We introduce random variables for our parameters, i.e., $\boldsymbol{P}$ ($\boldsymbol{P} = (P_1, \cdots, P_n)$) and $\hat{\boldsymbol{P}}$ ($\hat{\boldsymbol{P}} = (\hat{P}_1, \cdots, \hat{P}_n)$). The uncertainty of source trustworthiness may be due to lack of behavior consistency or experience, so the sources cannot provide stable-quality feedback [8, 22]. On the other hand, inadequate interaction with sources or inaccurate modeling by decision makers may incur uncertain trust estimation [10, 11]. Practical usage of WMV must have some procedures to arrive at values for $\hat{\boldsymbol{p}}$ to assign weights. Depending on the quality of the algorithm, there is a degree of correlation between trust and trustworthiness. We consider the procedure to get the unbiased trust: $\mathbf{E}(\boldsymbol{P}) = \hat{\boldsymbol{p}}$, which is a reasonable assumption for learning-based procedures [2, 13].

If the trustworthiness of a source is less than its trust, $p < \hat{p}$, then the actual correctness achieved is lower than what the decision maker believes: $\omega(\hat{\boldsymbol{p}}, \boldsymbol{p})$<$\omega(\hat{\boldsymbol{p}}, \hat{\boldsymbol{p}})$, while it is higher when $p > \hat{p}$. The main result is that, as long as the trust values used for decision-making are unbiased, the decision accuracy that $\mathcal{D}_W$ is believed to achieve $\omega(\hat{\boldsymbol{p}}, \hat{\boldsymbol{p}})$, equals what it actually achieves on average : $\mathbf{E}(\omega(\hat{\boldsymbol{p}}, \boldsymbol{P})) = \sum_{\boldsymbol{p}} \mathbf{Pr}(\boldsymbol{P} = \boldsymbol{p})\omega(\hat{\boldsymbol{p}}, \boldsymbol{p})$. We call this property *Stability of Correctness*, and prove it absolutely holds for WMV.

Theorem 1. *Stability of Correctness: If $\hat{\boldsymbol{p}} = \boldsymbol{E}(\boldsymbol{P})$, then*

$$\boldsymbol{E}(\omega(\hat{\boldsymbol{p}}, \boldsymbol{P})) - \omega(\hat{\boldsymbol{p}}, \hat{\boldsymbol{p}}) = 0 \tag{3}$$

A better procedure to obtain trust returns values closer to the trustworthiness values, with little variance. The quality of the procedure does not affect the stability of correctness at all when its unbiased on average, e.g., the shape of the trustworthiness distribution is irrelevant, which may initially seem counter-intuitive. However, *Stability of Optimality* captures the idea that even when its unbiased, poor trust values still result in worse performance of WMV. It measures the difference between the actual accuracy of decisions made using trust values $\mathbf{E}(\omega(\hat{\boldsymbol{p}}, \boldsymbol{P}))$, and of those made using trustworthiness values, i.e., $\mathbf{E}(\omega(\boldsymbol{P}, \boldsymbol{P})) = \sum_{\boldsymbol{p}} \mathbf{Pr}(\boldsymbol{P} = \boldsymbol{p})\omega(\boldsymbol{p}, \boldsymbol{p})$.

Theorem 2. *Stability of Optimality: If $\hat{\boldsymbol{p}}$=$\boldsymbol{E}(\boldsymbol{P})$ and all P_i have support $[\hat{p}_i - \delta_i, \hat{p}_i + \delta_i]$, then*

$$\boldsymbol{E}(\omega(\boldsymbol{P}, \boldsymbol{P})) - \boldsymbol{E}(\omega(\hat{\boldsymbol{p}}, \boldsymbol{P})) \leq \frac{1 - \omega(\hat{\boldsymbol{p}}, \hat{\boldsymbol{p}})}{2} \sum_{i=1}^{n} \frac{\delta_i}{1 - \hat{p}_i} \tag{4}$$

Although there is a gap between making decisions based on unbiased trust and based on trustworthiness, Theorem 2 proves that this gap is bounded by a relatively small threshold, implying that the unbiased trust would not reduce the decision quality too much. The upper bound is influenced by the distribution of trustworthiness, and converges towards zero with the variance of that reducing.

We also provide numerical experiments to demonstrate how the parameters influence Stability of Optimality, refer to our full paper [3]. The experiments imply that the accuracy gap is indeed sensitive to parameter δ, which depicts the range and variance of sources' trustworthiness, but counter-intuitively, is not sensitive to the number of sources.

4 DISCUSSION

For future work, more precise bounds for stability of optimality can be obtained with more detailed information provided, e.g., the variance of trustworthiness. Besides, it is worth investigating the stability of WMV in a more general case, namely when trust is a biased estimate of trustworthiness. It is a possible solution to distribute more estimate errors on sources that nearly have no influence on the decision result [1, 7]. Lastly, the proposed formal framework and the two types of stability can also be generalized to analyze the stability of other types of decision mechanisms.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation of China (NSFC) under Grants 62293510/62293511, and NSFC under Grant 62106223.

REFERENCES

[1] Tahar Allouche, Bruno Escoffier, Stefano Moretti, and Meltem Öztürk. 2021. Social ranking manipulability for the cp-majority, Banzhaf and lexicographic excellence solutions. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*. 17–23.

[2] Peter Auer. 2002. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research* 3, Nov (2002), 397–422.

[3] Shaojie Bai, Dongxia Wang, Tim Muller, Peng Cheng, and Jiming Chen. 2022. Stability of Weighted Majority Voting under Estimated Weights. *arXiv preprint arXiv:2207.06118* (2022).

[4] Daniel Berend and Aryeh Kontorovich. 2013. A sharp estimate of the binomial mean absolute deviation with applications. *Statistics & Probability Letters* 83, 4 (2013), 1254–1259. https://doi.org/10.1016/j.spl.2013.01.023

[5] Daniel Berend and Aryeh Kontorovich. 2014. Consistency of weighted majority votes. *Advances in Neural Information Processing Systems* 27 (2014).

[6] Daniel Berend and Aryeh Kontorovich. 2015. A finite sample analysis of the Naive Bayes classifier. *J. Mach. Learn. Res.* 16, 1 (2015), 1519–1545.

[7] Larry Bowen. 2009. Weighted voting systems.

[8] Alexander Philip Dawid and Allan M Skene. 1979. Maximum likelihood estimation of observer error-rates using the EM algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 28, 1 (1979), 20–28.

[9] David A Freedman. 1963. On the asymptotic behavior of Bayes' estimates in the discrete case. *The Annals of Mathematical Statistics* 34, 4 (1963), 1386–1403.

[10] Chao Gao, Yu Lu, and Dengyong Zhou. 2016. Exact exponent in optimal rates for crowdsourcing. In *International Conference on Machine Learning*. PMLR, 603–611.

[11] Pascal Germain, Alexandre Lacasse, Francois Laviolette, Mario March, and Jean-Francis Roy. 2015. Risk Bounds for the Majority Vote: From a PAC-Bayesian Analysis to a Learning Algorithm. *Journal of Machine Learning Research* 16, 26 (2015), 787–860.

[12] JQ James. 2020. Sybil attack identification for crowdsourced navigation: A self-supervised deep learning approach. *IEEE Transactions on Intelligent Transportation Systems* 22, 7 (2020), 4622–4634.

[13] Tor Lattimore and Csaba Szepesvári. 2020. *Bandit algorithms*. Cambridge University Press.

[14] Hongwei Li and Bin Yu. 2014. Error rate bounds and iterative weighted majority voting for crowdsourcing. *arXiv preprint arXiv:1411.4086* (2014).

[15] Edoardo Manino, Long Tran-Thanh, and Nicholas R Jennings. 2019. On the efficiency of data collection for multiple Naïve Bayes classifiers. *Artificial Intelligence* 275 (2019), 356–378.

[16] Tim Muller, Dongxia Wang, and Jun Sun. 2020. Provably Robust Decisions based on Potentially Malicious Sources of Information. In *2020 IEEE 33rd Computer Security Foundations Symposium (CSF)*. IEEE, 411–424.

[17] Shmuel Nitzan and Jacob Paroush. 1982. Optimal decision rules in uncertain dichotomous choice situations. *International Economic Review* (1982), 289–297.

[18] Theodoros Rekatsinas, Manas Joglekar, Hector Garcia-Molina, Aditya Parameswaran, and Christopher Ré. 2017. Slimfast: Guaranteed results for data fusion and source reliability. In *Proceedings of the 2017 ACM International Conference on Management of Data*. 1399–1414.

[19] Zhijun Tong and Richard Y Kain. 1991. Vote assignments in weighted voting mechanisms. *IEEE Trans. Comput.* 40, 05 (1991), 664–667.

[20] Elizabeth Walter. 2008. *Cambridge advanced learner's dictionary*. Cambridge university press.

[21] Bin Yu, Munindar P Singh, and Katia Sycara. 2004. Developing trust in large-scale peer-to-peer systems. In *IEEE First Symposium onMulti-Agent Security and Survivability, 2004*. IEEE, 1–10.

[22] Jing Zhang and Xindong Wu. 2018. Multi-label inference for crowdsourcing. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2738–2747.