Romanuke V.V.
Khmelnitskiy National University,
Khmelnitskiy, Ukraine,
E-mail: romanukevadimv@gmail.com

ACCURACY IMPROVEMENT IN WEAR STATE DISCONTINUOUS TRACKING MODEL REGARDING STATISTICAL DATA INACCURACIES AND SHIFTS WITH BOOSTING MINI-ENSEMBLE OF TWO-LAYER PERCEPTRONS

UDC 539.375.6+539.538+519.237.8

There is presented a method of improving accuracy in tracking metal tool wear states discontinuously, when the states' finite set has been statistically tied to the set of representative wear influencing factors. Range of wear states is presumed to be wholly sampled into those factors. The tracker is a static model based on boosting mini-ensemble of three two-layer perceptrons with nonlinear transfer functions. It regards statistical data inaccuracies and shifts. For making the ensemble, the AdaBoost technique is used. A distinction of the presented method of boosting from the AdaBoost is in the rule for finding the decreasing coefficient in order to re-distribute weights over training samples. Another one is that the ensemble is aggregated at once. The averaged gain of the boosting mini-ensemble in tracking 24 wear states with 16 influencing factors exceeds 50 %. The wear state tracking model is going to be perfected on optimizing two parameters of the training set and the naive rule for finding the decreasing coefficient before re-distributing training samples' weights.

Key words: wear state, statistical data, data jitter inaccuracies, data omissions, data shifts, tracking model, accuracy, two-layer perceptron, boosting, boosting ensemble, tracking error rate.

Benefits of tracking wear states discontinuously

In industrial metal processing, rationalized usage of billets and tools is desired. This is partially realized with tracking metal wear states letting prevent underuse and overuse. Discontinuousness of the tracking benefits because its reliability and accuracy are far beyond higher than for continuous approach needing corrections of coefficients in differential and difference equations as time goes by. Besides, decision making on the usage is always of a discrete set of wear states, whose tracking accuracy is never perfect.

Approaches to wear state tracking accuracy improvement

Improvement of the wear state tracking accuracy is predetermined with the way of tracking wear states. While tracking discontinuously, a finite set of wear states is controlled by a finite difference method or a method of statistical correspondence [1]. The finite difference method accuracy is increased only with more accurate sampling. And the accuracy rank of statistical approximation is dependent on the initial statistical data. With statistically universal approximators based on two-layer perceptron with nonlinear transfer functions (2LPNLTF), it is possible to make the rank higher using specific training samples which could regard statistical data inaccuracies and shifts [2]. Based on boosting technique [3, 4], ensembles of weak learners may increase the accuracy rank additionally. However, when the number of wear influencing factors (WIF) is of the order of tens [5], such ensembles requiring a few 2LPNLTF haven't been tried on their performance. In particular, AdaBoost technique (ABT) uses just learners of small configuration [6, 7], although 2LPNLTF might be tried as well.

The article goal

Under supposition that there is a few tens of WIF corresponding to each known wear state within a finite statistical data set (FSDS), we are to increase the wear state 2LPNLTF-tracking accuracy with mini-ensemble of 2LPNLTF. The ensemble performance is going to be boosted on the basis of ABT. Before boosting, the 2LPNLTF classifier along with its training routine is formalized. The real gain of the boosting will be discussed to focus, probably, on wear state tracking problems having different numbers of WIF, wear states, and engaging various numbers of 2LPNLTF in ensembles.

Tracking wear states with boosting mini-ensemble of 2LPNLTF

The single object input of 2LPNLTF is $\mathbf{X} = \left[x_i \right]_{1 \times Q} \in X \subset \mathbb{R}^Q$ with $Q \in \mathbb{N}$ WIF and the output of 2LPNLTF is the number [2]

$$s_* \in \arg\max_{s=1,\,N} \left\{ \left(1 + \exp\left[-\left(\sum_{k=1}^{S_{\mathrm{SHL}}} u_{ks} \cdot \left(1 + \exp\left[-\left(\sum_{i=1}^{Q} x_i a_{ik} + h_k \right) \right] \right)^{-1} + b_s \right) \right] \right)^{-1} \right\} = \arg\max_{s=1,\,N} v_s \qquad (1)$$

of the current wear state for the number $N \in \mathbb{N} \setminus \{1\}$ of total states, where S_{SHL} is number of neurons in the single hidden layer of 2LPNLTF, and $S_{\text{SHL}} \cdot (Q+N+1)+N$ coefficients

$$\left\{ [a_{ik}]_{Q \times S_{\text{SHL}}},\, [u_{ks}]_{S_{\text{SHL}} \times N},\, [h_k]_{1 \times S_{\text{SHL}}},\, [b_s]_{1 \times N} \right\} \tag{2}$$

are to be determined in the training process. FSDS is $\left\{ \mathbf{X}_j, w_j \right\}_{j=1}^{L}$ by $L \in \mathbb{N} \setminus \left\{ \overline{1, N-1} \right\}$ and $\mathbf{X}_j = \left[x_i^{\langle j \rangle} \right]_{1 \times Q} \in X$ corresponding to the wear state $w_j \in \left\{ \overline{1, N} \right\}$. All possible wear states are represented in FSDS: $\left\{ w_j \right\}_{j=1}^{L} \cap \left\{ \overline{1, N} \right\} = \left\{ \overline{1, N} \right\}$. Henceforward, the s-th state w_{j_s} is reflected with the pure representative $\mathbf{X}_{j_s}$. The input of 2LPNLTF is fed successively with the training set

$$\left\{ \mathbf{Y} = [y_{is}]_{Q \times N} : y_{is} = x_i^{\langle j_s \rangle} \right\} \tag{3}$$

and
$$\left\{ \left\langle \{\mathbf{Y}\}_{r=1}^{R}, \{\tilde{\mathbf{Y}}_h\}_{h=1}^{H} \right\rangle : \tilde{\mathbf{Y}}_h = \mathbf{Y} + \sigma_h \cdot \mathbf{\Xi} + \mu \cdot \sigma_h \cdot \mathbf{\Theta},\, \sigma_h = h \sigma_0 H^{-1}\ \forall h = \overline{1, H},\, H \in \mathbb{N}, \right.$$

$$\left. \sigma_0 > 0,\, \mathbf{\Xi} = [\xi_{is}]_{Q \times N},\, \xi_{is} \in \mathrm{N}(0, 1),\, \mu > 0,\, \mathbf{\Theta} = [\theta_{is}]_{Q \times N},\, \theta_{is} = \zeta_s \in \mathrm{N}(0, 1) \right\} \tag{4}$$

by $R \in \mathbb{N} \cup \{0\}$ and the infinite set $\mathrm{N}(0, 1)$ of standard normal variate's values. For making sure that the pure representatives (3) have not been disassociated from those N wear states, the input of 2LPNLTF is re-fed with the set (3). Having got $Q = 16$, $N = 24$, $S_{\text{SHL}} = 45$, $\left\{ x_i^{\langle j_s \rangle} \in [0; 1] \right\}_{i=1}^{16}$ in a problem of 24 wear states tracking, 2LPNLTF (1) was identified by three methods of determining coefficients (2) — "traingda" ($\alpha = 1$), "traingdx" ($\alpha = 2$), "trainscg" ($\alpha = 3$). FSDS for ABT is formed via the set (4) by $R = 1$, $H = 20$, $\sigma_0 = 0.25$, $\mu = 1.5$, whereupon this set $\mathbf{P}_t$ is re-generated for $T = 100$ times. Thus FSDS for ABT is $\{\mathbf{P}_t\}_{t=1}^{100}$ including

$$(R+H) \cdot N \cdot T = (1+20) \cdot 24 \cdot 100 = 50400$$

training samples. At the q-th iteration of boosting, these samples have the weights in $\mathbf{D}(q) = \left[d_\tau(q) \right]_{1 \times 50400}$ by $q = \overline{1, q_0}$ at some final iteration number q_0.

Initially, $d_\tau(1) = 50400^{-1}\ \ \forall \tau = \overline{1, 50400}$. Matrix $\mathbf{A} = [\bar{a}_{\alpha\tau}]_{3 \times 50400}$ is of correct responses of classifiers, where $\bar{a}_{\alpha\tau} = 1$ is the correct classification of τ-th sample by the α-th 2LPNLTF, otherwise $\bar{a}_{\alpha\tau} = 0$. The weighted errors are in the matrix $\mathbf{E}(q) = \left[\eta_\alpha(q) \right]_{3 \times 50400}$, where

$$\eta_\alpha(q) = \sum_{\tau=1}^{50400} d_\tau(q) \cdot (1 - \bar{a}_{\alpha\tau}),\ \alpha = \overline{1, 3}. \tag{5}$$

Starting from $q = 1$, there are found each classifier's weighted error (5), the best 2LPNLTF

$$\alpha_*(q) \in \arg \min_{\alpha = \overline{1, 3}} \eta_\alpha(q), \tag{6}$$

and the minimal weighted error

$$\eta_*(q) = \min_{\alpha = \overline{1, 3}} \eta_\alpha(q), \tag{7}$$

letting learn the coefficient

$$\gamma(q) = 1 - \eta_*(q) \tag{8}$$

and calculate the new distribution $\mathbf{D}(q+1)$ of weights

$$d_\tau(q+1) = \tilde{d}_\tau \Big/ \sum_{\upsilon=1}^{50400} \tilde{d}_\upsilon \ \text{ by } \ \tilde{d}_\tau = d_\tau(q) \cdot \exp\left[-\gamma(q) \left(2 \cdot \bar{a}_{\alpha_*(q), \tau} - 1 \right) \right] \tag{9}$$

over samples in $\{\mathbf{P}_t\}_{t=1}^{100}$. If $\eta_*(q) < 1 - N^{-1}$ then $\tilde{q} = q$ and $q = \tilde{q} + 1$, and (5) — (9) are re-found. If $\eta_*(q) \geq 1 - N^{-1}$ then $q_0 = q$ and there are calculated the following coefficients:

$$\tilde{\gamma}(q)=\gamma(q)\Big/\sum_{p=1}^{q_0}\gamma(p) \text{ by } q=\overline{1,q_0} \text{ for } \beta(\alpha)=\sum_{q\in\left\{\overline{1,q_0}\right\},\ \alpha=\alpha_*(q)}\tilde{\gamma}(q). \quad (10)$$

Denoting the α -th 2LPNLTF value v_s as $v_s(\alpha)$, the boosted classifier output is

$$s_*\in\arg\max_{s=\overline{1,24}}\tilde{v}_s \text{ by } \tilde{v}_s=\beta(1)v_s(1)+\beta(2)v_s(2)+\beta(3)v_s(3). \quad (11)$$

In the section below we'll see the real gain of the boosting mini-ensemble of three 2LPNLTF. The gain is defined with the tracking error rate (TER) being the percentage of the classifier's correct responses among the total inputs. Particularly, ratio between TER of a 2LPNLTF and TER of the ensemble in (11) is of interest.

Results and discussion

It has been exposed experimentally that TER of the ensemble in (11) is about 56 % lower than TER of the best single 2LPNLTF in tracking 24 wear states. At that the worst ensemble TER is no greater than 1.12 %, and 0.96 % is TER of the best-combined ensemble. At the highest level of jitter inaccuracies and omissions in statistical data or measurements and WIF shift in every state, the mean ensemble TER is 6.82 %, although the gain of the boosting is only 39 %. Another side of the boosting mini-ensemble effectiveness is that variance of wear states' TER is more than 50 % lower. These results are evidence of the noticeable gain of the boosting based on just three 2LPNLTF. Nearly the same effects were locally registered on some other wear state tracking problems. Consequently, the boosting gain is expected also in solving problems having different numbers of WIF or (and) wear states. Very likely, engaging four and more 2LPNLTF in ensembles will force the gain, improving accuracy further.

A distinction of the presented method of boosting from ABT is in the coefficient (8). Another one is that the ensemble is aggregated at once. It is fully realizable within MATLAB, and its code may be freely edited for adapting to specified problems of wear state tracking. And as the tracker is ensemble of 2LPNLTF then the classifier operation speed is high enough. While MATLAB-tracking on Intel Core i3-4150 CPU @ 3.50 GHz with 4 GB RAM within 64-bit Windows 7, the ensemble tracks over 2300 wear states per second. Importantly, that the improved averaged accuracy of the ensemble is maintained constant through the whole range of wear states.

Conclusion

The boosting mini-ensemble of three 2LPNLTF appears capable to track wear regarding statistical data inaccuracies and WIF shifts more accurately. By this model, some WIF at the input may be even omitted. Omissions are substituted with elements from $\mathrm{N}(0,1)$. However, the accuracy might be improved more, because parameters $T=100$ and $H=20$ were assigned empirically. They could be optimized. Also the naive rule (8) is possibly nonoptimal for the decreasing coefficient $\gamma(q)$ in re-distributing weights (9). Hence, it is well-promising that the wear state tracking model is going to be perfected.

References

1. Chungchoo C. On-line tool wear estimation in CNC turning operations using fuzzy neural network model / C. Chungchoo, D. Saini // International Journal of Machine Tools and Manufacture. – 2002. – Volume 42, Issue 1. – P. 29-40.

2. Romanuke V. V. Wear state discontinuous tracking model as two-layer perceptron with nonlinear transfer functions being trained on an extended general totality regarding statistical data inaccuracies and shifts / V. V. Romanuke // Problems of tribology. – 2014. – N. 3. – P. 50-53.

3. Champion M. Sparse regression and support recovery with image-Boosting algorithms / M. Champion, C. Cierco-Ayrolles, S. Gadat, M. Vignes // Journal of Statistical Planning and Inference. – 2014. – Volume 155. – P. 19-41.

4. Shen C. Fully corrective boosting with arbitrary loss and regularization / C. Shen, H. Li, A. van den Hengel // Neural Networks. – 2013. – Volume 48. – P. 44-58.

5. Hua J. A cobalt diffusion based model for predicting crater wear of carbide tools in machining titanium alloys / J. Hua, R. Shivpuri // Journal of Engineering Materials and Technology. – 2005. – Volume 127. – P. 136-144.

6. Nie Q. Probability estimation for multi-class classification using AdaBoost / Q. Nie, L. Jin, S. Fei // Pattern Recognition. – 2014. – Volume 47, Issue 12. – P. 3931-3940.

7. Sun Y. Unifying multi-class AdaBoost algorithms with binary base learners under the margin framework / Y. Sun, S. Todorovic, J. Li // Pattern Recognition Letters. – 2007. – Volume 28, Issue 5. – P. 631-643.

Романюк В. В. **Покращення точності у дискретній моделі відслідковування стану зносу з урахуванням похибок і зсувів у статистичних даних на основі міні-комітету бустингу двошарових персептронів.**

Представляється метод покращення точності дискретного відслідковування станів зносу металевого засобу, коли скінченна множина цих станів була статистично пов'язана з множиною репрезентативних факторів, що впливають на знос. Діапазон станів зносу вважається повністю розбитим за цими факторами. Відстежувачем є статична модель на основі міні-комітету бустингу трьох двошарових персептронів з нелінійними передавальними функціями. Вона враховує похибки і зсуви у статистичних даних. Для утворення згаданого комітету використовується техніка адаптивного бустингу. Одна з відмінностей методу бустингу, що представляється, від адаптивного бустингу полягає у правилі знаходження спадного коефіцієнта для того, щоб перерозподіляти ваги навчальних зразків. Ще одна полягає у тому, що комітет утворюється одразу. Усереднений виграш такого міні-комітету бустингу у відслідковуванні 24 станів зносу з 16 факторами впливу перевищує 50 %. Дана модель відслідковування стану зносу буде удосконалена завдяки оптимізації двох параметрів навчальної множини і наївного правила знаходження спадного коефіцієнта перед перерозподілом ваг навчальних зразків.

Ключові слова: стан зносу, статистичні дані, флуктуаційні похибки у даних, пропуски у даних, зсуви у даних, модель відслідковування, точність, двошаровий персептрон, бустинг, комітет бустингу, рівень помилок відслідковування.