High Accuracy Dynamic Time Warping with Many-to-Many Technique to Classify Abnormalities of Free Range Cattle

Rattapon Dulyala

Department of Engineering, Mahasarakham University, Mahasarakham, Thailand.

e-mail: rattapon.dul@uru.ac.th; apinan.a@msu.ac.th

***Abstract* - Detection of free range husbandry's cow abnormalities is considered as a first priority for farmer. However, it is difficult to observe and classify these abnormalities all time. Then, leads to some important issue such as the artificial insemination problems and problem of productivity decrease. This paper proposed high accuracy Dynamic Time Warping with many-to-many technique for classify abnormalities of free range cows due to in original version of Dynamic Time Warping still has a straight-line adjustment problems of multiple point on one time series to one point on another time series, which may result in incorrect classification when used to classify time series. A three-axis accelerometer was used to collected behavior of walking-grazing behavior, standing behavior and lying behavior of five cows in 30 days, dataset from first 10 days was used as a value of referential and dataset from last 20 days used in comparing and measuring the similarities with the referential. The result show 100% accuracy in detecting abnormalities due to leg injuries when compare with expertise farmers or veterinarians.**

Keywords – component, Dynamic Time Warping, Many-to-Many technique, High accuracy, Free rang cattle

I. INTRODUCTION

Dynamic Time Warping (DTW) is one of the most widely used gauges, particularly in measuring the distance between the time series for use in classification or clustering including prediction [1]. The reasons why DTW is popular because of it is an elastic measurement that can work efficiently even with the shift of data on the axis of time. By this attribute, DTW is gaining attention in many areas of research and development. However, DTW calculate by uses only the values of each data point of both series, without in warped paths of multiple points on one time series to one point on another time series or singularity and may result in incorrect classification when leading used to classify time series. In the past few years, this problem received the attention of researchers in presenting research results that can solve this problem. One of them is the presentation of a weighted DTW [2] that can solve these problems very well, but instead comes with additional parameters that are difficult to find the right value.

This research proposes a measure of distance that still has good characteristics of the traditional DTW, while can solve the warped path of the traditional DTW. By adding a process of comparing the values of a series to a thorough, many-to-many, which gives a more varied and detailed reference point. When applied to classification or prediction will be more accurate than using traditional DTW. Then, this algorithm is used to identify abnormalities of free range cows by measuring the similarity of the duration of their behaviors and compare with individual dataset from sensor which install with front leg of free range cows. Then, summarize the abnormalities for rancher to check their cows health and evaluate the accuracy of proposed algorithm.

II. LITERATURE REVIES

Many researchers have applied and developed automatic systems to sort their abnormalities with embedded systems stalled on their legs or necks to measure the duration of their daily behaviors. This approach has been used for sorting lameness [3]. While [4-6] used statistical tests where a mean of recent activity was compared to hind cast mean values. Analyses of time series in which linear Kalman filters estimated parameters, were presented in [5]. Furthermore, [7] detected estrus by Fuzzy logic methods. A Generalized Likelihood Ratio (GLR) test was adopted to observe distributions of activity data by [8] and further improved in [9] Fuzzy Logic classification of the alerts utilizing the period between estruses. Although these studies were correct and precise, they required a significant amount of data for analysis and using complicated analysis methods.

III. MATERIALS AND METHOD

This research proposed to use Many-to-Many technique with Dynamic Time Wraping to improve problem of straight-line adjustment problems of multiple point on one time series to one point on another time series . Then, try to evaluated its performance with experiment in order to classifying the abnormalities of the free range cows we using the data set from sensors that attached to cows as examples all detail can described as below topic.

A. Many-to-many technique

Many-to-many technique is refer from one type of relationship between entities in database management, that

explained as one or more records in the first entity that have relationship with members of the second entity in multiple records. At the same time, if looking back, members in the second entity return to have relationships with members of the first entity in the amount of many records as well. By this way, we then tried to apply this pattern to the matching of time series in dynamic time warping because of in general, dynamic time warping is a method that allows computers to find the appropriate pairing of two sets under the constraint. Those sequences are distorted wrap in units of time. In order to consider the similarity from the unstable distribution in time units. Which will give the result the distance and the best alignment, but in reality, dynamic time warping still has a straight-line adjustment problems of multiple point on one time series to one point on another time series. Which may result in incorrect classification when used to classify time series. With the method of increasing even distribution, this problem can be reduced.

B. Dynamic Time Wraping

Dynamic Time Warping (DTW) is the dynamic programming for comparing similarities between 2 sets of time series. The comparing result will be determined the most reliable data between distance value and alignment which is flexible and well-supported in case of an occurrence in time axis variation [12].

For example, comparing the similarity of X time series with Y time series where the length of data was symbolized as N and M respectively.

$x_1, x_2, \ldots, x_i, \ldots x_N$

$y_1, y_2, \ldots, y_i, \ldots y_M$

Whereas,

X is the time series $x_1, x_2, \ldots, x_i, \ldots x_N$

Y is the time series $y_1, y_2, \ldots, y_i, \ldots y_M$

The 2 matrixes were created based on both the time series of data which the X matrix has the equal number of rows as N data columns were copied starting from left to right and having the same size. For X matrix, every row must be equally copied starting from the upper to the bottom of dimension of X matrix. After obtaining X matrix and Y matrix, the absolute difference of both matrixes was calculated based on equation [13-15]. The result was D matrix:

$$d(i,j) = \sqrt{(x_i - y_i)^2} \qquad (1)$$

Cost matrix was created by each element of accumulated cost matrix using equation below:

$$D(i,j) = d(i,j) + \min \begin{cases} d(x_{i-1}, y_{i-1}) \\ d(x_i, y_{i-1}) \\ d(x_{i-1}, y_i) \end{cases} \qquad (2)$$

The calculation's result revealed that the wrap path is the shortest path of the cost matrix starting from D(N,M) to (1,1) by using the principle of similarity calculation of data in elements of matrix from the horizontal line, vertical line and crossing line that could be conducted by choosing the shorted path.

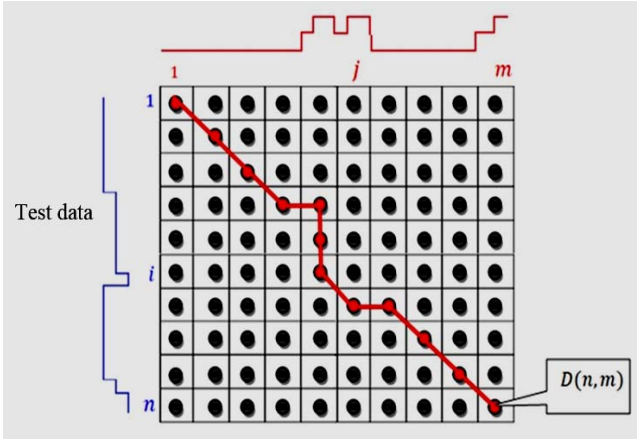

Figure 1. Accumulated cost matrix and optimal warping path.

C. Animals used in the experiment

The experiment animals include 55 American Brahman cows which are considered as popular species in Thailand. All cows aged 1-2 years, these 55 male cows belong to the villagers in Uttaradit Province, which located at north part of Thailand. All of them are free rang cows, released freely during 8.00 AM. – 6.00 PM. The accurate measurement of the abnormality classification process and the referential time series of normal cows' behaviors were tested by their data collected for 10 days (1^{st}-10^{th} days: as referential) and the daily results of the referential values was individually revealed.

D. Experiment devices and installation

The 3-axis accelerometer using MODULE ET-MMA7331L from previous study [11] was employed to each cows. The gravity was set at ±4g. The value from the sensor refers to pressure that transmits data to microprocessor Atmega328. Then this value was converted into gravity (g). The device was powered by Lithium Polymer battery (Li-Po) 3.7V 500 mAh. An accelerometer was attached on the top of the front leg of each individual cows. The values measured on the three axes was imported to a wireless transmitter. A series of data was transmitted in every one second (1s/1) by Xbee pro series 2. This device transmitting data in the 2.4

GHz. Frequency was placed on the front leg of each individual cows along with the transceiver using module called Bluebee Dongle which displays behavior status of the cows at different periods of time. The transmitted data displayed behavior status in a program named real-time monitoring behavior of free range cows. This program was designed to display the value from the sensor, total acceleration, and of physical behaviors such grazing behavior, standing behavior and lying behavior.

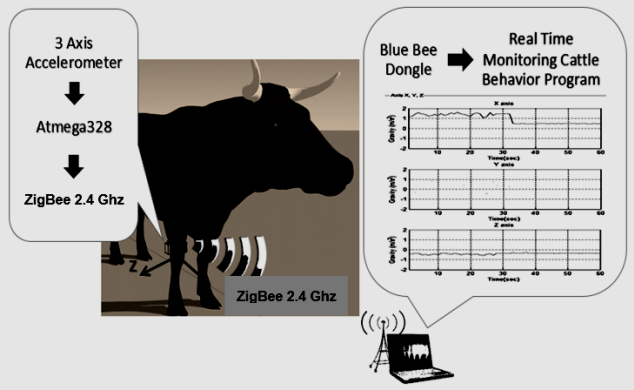

Figure 2. Device installation and accelerometer sensor.

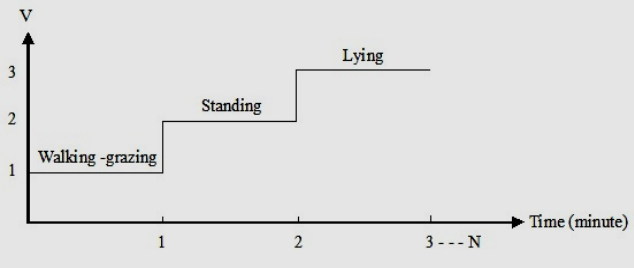

Figure 3. Time Series of Cow's behavior.

After this step, we sorting of cows abnormalities by measuring the similarity of the duration of their behaviors through Dynamic Time Warping with many-to-many technique. The sorting process was divided into two parts: 1) calculating referential similarity based on the behavior of all cows in normal conditions that was collected as a database in the form of time series data. All data was matched and compared the similarity. The obtained result was the group of referential accumulated distance and the maximum reference accumulated distance was used as the criteria for sorting the behavior of cows. While 2) Time series of the current behavior of the cows were used for comparing similarity with time series of each database. The result was a group of accumulated distance. If this accumulated distance were higher than the referential accumulated distance, it would indicated that the cow had some abnormalities.

IV. RESULS AND DISCUSSIONS

The time series data of free range cows were collected as a referential database for sorting cows health condition. The referential value of a singles cow could be found by using 10 days of time series to compare through paired-comparison and similarity measured using dynamic time warping.

A. Referential Accumulated Distance of a Single Cow

The time series data of 55 cows were collected as a referential database for sorting cow health condition. The referential value of a single cow could be found by using 10 days of the time series to compare through paired-comparison and similarity measured using dynamic time warping.

The time series of the cow data X(N+1) was brought to the sorting process by measuring the similarity of a time series of the referential database. The result appeared accumulated scores of N sets. These sets was brought to the decision making process with 3 impossible cases as shown in figure 4.

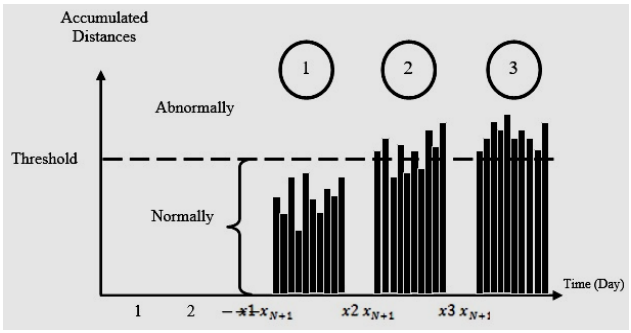

Figure 4. Health sorting of cows.

The referential value of a single cow could be found by using 10 days of the time series to compare through paired-comparison and similarity measured using dynamic time warping. Figure 5 represents 45 pairs of the accumulated distance of 1st cow and the accumulated distance of the pair of the 2nd and 9th days had the highest value and in figure 6 represents accumulated matrixes and wrap path by using MATLAB program in which the maximum value accumulated distance was 31. Among 45 pairs distributed form 0-31, the value identified the similarity of referential time series data whereas the maximum accumulated distance was used as a threshold for sorting cows health.

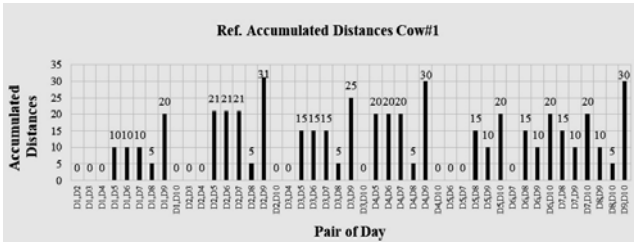

Figure 5. Total accumulated distance of referential time series database

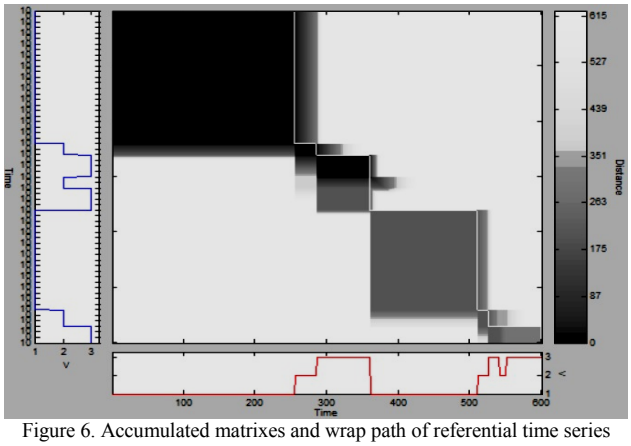

Figure 6. Accumulated matrixes and wrap path of referential time series

B. Classification of Health Base on Dynamic Time Wrapping with Many-to-Many Technique.

According to cows health sorting based on dynamic time warping by using current time series data to measure the similarity of reference time series form the 1st - 10th days, the obtained result was the group of accumulated distance tested. The sets of accumulated distance based on the threshold value if they were lower than the referential accumulated distance, it indicated that cows has normal behaviors. Therefore, the data were stored for conducting the paired-comparison with referential time series. However, if there was no sign of any abnormalities. The data of every single day were continually stored until there was an abnormality of the cow.

In the case of the 1st cow, the accumulated distance of all 10 datasets was higher than referential accumulated distance. This result revealed that accumulated distance of all data sets was higher than maximum the referential accumulated distance, resulting that they were abnormal shown in figure 5. Observation of the 1st cow behaviors on the 16th day indicated that it slept longer than other days because it was sick. However, on the 20th day, it was sick (has a fever).

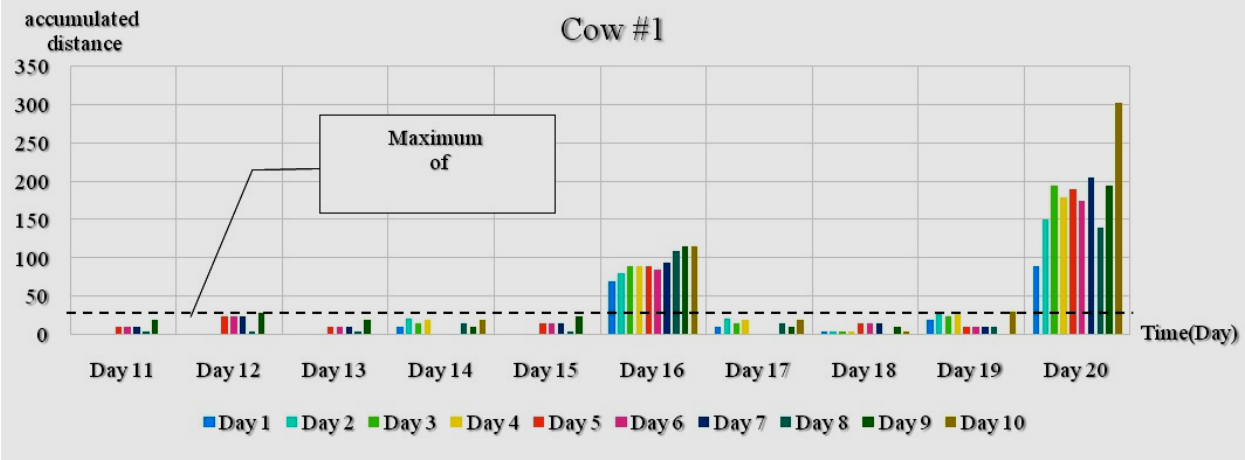

Figure 5. Example comparison of accumulated distance from the 11th -20th days with referential accumulated distance that was in grey tab.

V. CONCLUSIONS AND FUTURE WORK

The classification of the cow abnormalities is particularly based on the time series of the current cows data that was brought into the sorting process, the similarity measurement and time series of referential database, the accumulated score sets were revealingly formed as an indicator of cow conditions. This indicator was used for their health classification. As the result of this process, 8 cows were found some abnormalities because their accumulated distances were over the threshold level. Meanwhile other 47 cow's observation of its accumulated distances was lower than the threshold level which indicated that the cow was in the normal condition. When compare our results with farmer's observation and health test in 8 cows by veterinarians that our proposed algorithm indicate "abnormal", the results show that all classified are correct. Therefore, accuracy can be calculated at 100%.

Based on the evaluation results of purposed method, the authors found that the possible reasons for better performance caused by

1) Tools (sensor) for collecting data of free range cows behavior are suitable for actual use and has better performance than even use before in [11] by adjust some parameter. Thus, makes the behavior data collection more detailed and accurate.

2) Matching algorithm using many-to-many technique can compare the similarity of two dataset more completely [16]. Resulting in better predictive results with dynamic time warping.

Nevertheless, the referential time series of other various behaviors should be investigated for more effective dynamic time warping.

ACKNOWLEDGMENT

The authors so deeply impressive and wish to thanks colleagues in Uttaradit Rajabhat University for all aspects of support and assistance throughout the time that the author has to collect data in fieldwork until the experiments is completed. Next, wish to thanks faculty of Engineering, Mahasarakham University for all support tools and helpful advice. In addition, the authors wish to thanks the Associate Editor and all referees for their many helpful comments and suggestions.

REFERENCES

[1] X.Ying and W.Bryan. "Adaptive feature based dynamic time warping" International journal of computer science and network security. Vol. 10, 2010, pp.264-273.

[2] Y.S. Jeong, M.K,Jeong and O.A.Omitaomu. " Weighted dynamic time warping for time series classification" Pattern recognition, Vol.44, 2011,pp. 2231-2240.

[3] A.Aurosopon and S.Kuankid. "Dynamic time wraping for classifying lameness in cows". Applied animal behaviour science, Vol. 12, 2016, pp. 134-141.

[4] L.A.Gonzalez, B.J. Tolkamp, M.P.Coffey, A.Ferret, I.Kyriazakis. "Changes in feeding behavior as possible indicators for the automatic monitoring of health disorders in dairy cows" Journal of dairy science, Vol.91, 2008, pp.1017-1028.

[5] N.Blackie, E.Bleach, J.Amory, J.Scaife. "Impact of lameness on gait characteristics and lying behaviour of zero grazed dairy cattle in early lactation" Applied animal behaviour science, Vol.129, 2011, pp.67-73.

[6] A.Shokoufandeh, Y.Keselman, M.F.Demirci, D.Macrini, S.Dickinson. "Manay-to-many feature maching in object recognition: a review of three approaches" IET Computer Vision, Vol.6, Iss.6, 2012, pp.500-518.

[7] H.A.Zarchi, R.I.Jonsson, M.Blanke. "Improving oestrus detection in dairay cows by combining statistical detection with fuzzy logic clsssification" Sci.agric, Vol.67, No.5,2010.

[8] M.Pastell, J.Tiusanen, M.Hakojarvi and L.Hannien. "A wireless accelerometer system with wavelet analysis for assessing lameness in cattle." Biosystems engineering journal, Vol.104,2009, pp. 545-551.

[9] P.Martiskainen, M.Jarvinen, J.P.Skon, J.Tiirikainen, M.Kolehmainen and J.Mononen. "Cow behavior pattern recognition using a three-dimensional accelerometer and support vector machines" Applied animal behavior science, Vol.119, 2009, pp. 32-38.

[10] A.Helenahokkanen, L.Hamminen, J.Tiusanen and M.Pastell. "Predicting sleeping and lying time of calves with a support vector machine classifier using accelerometer data" Applied animal behavior science, Vol. 34, 2011,pp. 10-15.

[11] R.Dulyala, S.Kuankid, T.Rattanawong, A.Aurasopon. " Classification system for estrus behavior of cow using an accelerometer", Asia-Pacific signal and information processing association annual summit and conference, APSIPA 2014, pp.

[12] K.Moriya, T.Yoshimura, M.Kitagawa, M.Oyamada and Y.Sugimoto. "Monitoring cattle behavior in the forest using GPS records" Nihon chikudangakkaiho, Vol. 74, No.2, 2003, pp.229-234.

[13] B.Robert, B.J.White, D.G.Renter and R.L.Larson. "Evaluation of three-dimensional accelerometer to monitor and classify behavior pattern in cattle" Computers and electronics in agriculture, Vol.67, 2009, pp.80-84.

[14] J.Zhao, Y.Liu, K.K.Chai, Y.Chen, and M.Elkashlan. "Many-to-many maching with externalities for device-to-device cmommunications". IEEE wireless communications letters, Vol.6, No.1, 2017, pp. 138-141.

[15] J.Grosso, A.Jhumka, M.Bradbury. "Reliable many-to-many routing in wireless sensor networks using ant colony optimisation", 15th European dependable computing conference (EDCC), 2019, pp. 111-118.

[16] S.Saha, O.Landsiedel and M.C.Chan. "Effeicient many-to-many data sharing using synchronous transmission and TDMA" 13th International conference on distributed computing in sensor systems, 2017, pp.19-26.