Badminton player scouting analysis using Frequent Pattern growth (FP-growth) algorithm

L Ardiantoro[1] and N Sunarmi[2]

[1] Informatics Engineering, Universitas Islam Majapahit, Jl. Raya Jabon, Mojokerto, Indonesia

[2] FTIK STAIN Tulung Agung, Jl. Mayor Sujadi 46, Tulungagung, Indonesia

E-mail: ipan.ardianto@gmail.com

Abstract. FP Growth algorithm is widely used to analyze patterns from a huge amount of data with (frequent) repeated items. The objective of this research is to analyze playing pattern a badminton player, one of a popular sport in Indonesia. The data set was generated from a technical stroke during the game. The model used in this study was Jonathan Christie a top Indonesian badminton player. The method of data collection was done by dividing the playing field into various areas of the game. Observations were made by using the software, to calculate and classify the types of stroke that carried out by the athlete. The result of this research; the tactical approach of Jonathan Christie during this match was described. The data obtained would be very useful for the coach to improve the athlete's performance. Another advantage obtained was the analysis of the athlete's performance can be done with a quantitative approach so that it can enrich the current methods. As the conclusion, the FP Growth algorithms were able to describe the game pattern of a badminton athlete, JC by using PHP and MySQL. Sport science has become a necessity to develop to increase athletes' competitiveness.

1. Introduction

Badminton is a popular sport in Indonesia and has produced many proud achievements in this country. Therefore, Indonesia needs to make sustainability effort, in order to remain competitive in major tournaments for these following years. Sport science implementation is essential for this level. One of them is implemented in the scouting process. The process can be done before, during and after the match in order to boost the athlete's achievements. A series of analysis is carried out, thus providing an exact approach in understanding the strengths and weaknesses, or opportunities of the opponent/athlete. The development of each player is monitored with meticulous data, and then formulated in an effective training program to boost performance [1]. Match preparation is a critical process for the highest level of performance. Through the development of systematic training plans that draw through upon knowledge garnered from a vast array of scientific discipline. The process of training and athlete performance is correlated with the execution of these various tasks in a high level of perfection and details. A scout is a person sent out to obtain information about players by watching them in action with a view to making recommendations.[2] We developed a supporting program to capture the data during the match. The software developed using PHP and MySQL. By the end of the match, scout could give an advice to coach the team for the next upcoming game. We focus this research in statistical data analysis for Indonesian player, Jonathan Christie.

2. **Literature review**

2.1. Badminton

Badminton is a sport using a racket and a shuttlecock that is struck to pass through a net so that the shuttlecock falls in the opponent's playing field [3]. The stroke techniques are used in this game are service, lob, smash, drop shot, and netting. Base on the direction part of the stroke, there are known as backhand and forehand, instead.

2.2. FP growth algorithm

The process of looking information from a population of selected amount data using an algorithm was called knowledge data discovery in data mining. FP-Growth is one of the a-priori algorithms in data mining, to find the rules of association between item combinations by considering the frequency of the data. [4]. The FP-Growth algorithm is an algorithm used in pattern selection to speed up the decision process frequent itemset before generating rule as a recommending decision. The apriori executions rule algorithm or association analysis is a data mining technique to find the rule of association between item combinations. This hidden relationship can be represented in the form of association rules for the set of items that often appear [5] [6]. Two parameters that considered in this algorithm are: support and confidence. Support means the combinational item from database, hence confidence is a strong relationship between items in associative rule. These parameters are formulated as follows [7] [6] :

$$Support\ (X) = \frac{Freq.\,(X,Y)}{Total\ Transactions\ (N)}$$

.... (1)

Confidence is a measure of how often item Y appears in transactions containing item X. Confidence is calculated after the support is determined. The confidence value of rule A given B is obtained from the following formula.

$$Confidence\ (Y|X) = \frac{Freq.\,(X,Y)}{Freq.\,(X)}$$

………………..(2)

The FP-Growth algorithm needs to generate candidate to obtain a frequent item set, by constructing a tree to seeking frequent item occurrences, called FP-Tree. This characteristic made the algorithm more efficient to structure the data to compare to other a-priori algorithms. [8]. By using a FP-Tree mechanism, FP-Growth has an ability to extract a frequent item set from data set directly. There are three steps as follows.[9] :

      1). Building conditional pattern base
      2). Generate conditional pattern base
      3). Filter and selection of Frequent item set

The first stage is collecting all item set that meets the support requirements for ≥ minimum support. These item sets are called undefined item set. The second stage is rule generation, aimed at forming rules with high confidence items from an unspecified item set that have been obtained previously. This rule is called strong rules. [7] [10]

3. **Methodology**

First of all, we do a literature review of badminton theory, tactics and rule of the game. Then we developed model that suitable with the needed data set by FP-Growth. The data model was developed by categorizing based on badminton field areas that most likely used by a player to do a specific technical stroke (Fig.1). After several times of service exchanges during the game, usually an athlete's pattern has been formed. So far this can be known based on subjective analysis form the observations of a coach who

is very observant only. By using such developed computer based program, this analysis can be done more objectively based on statistical data that is easier to analyze and archive in dBase form. On the other hand, this computer program can also be used to analyze the game of a post-match analysis based on video file(s).
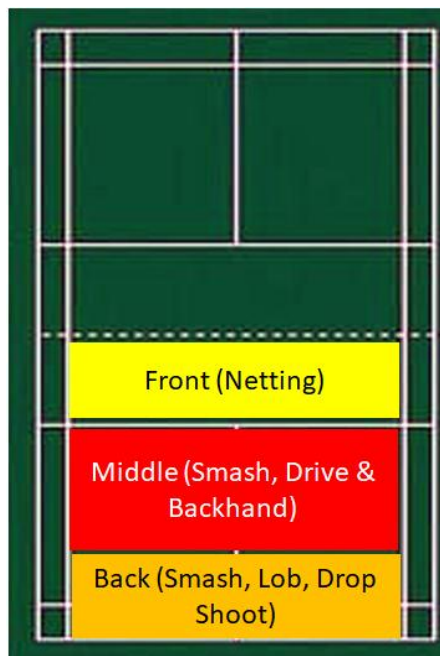
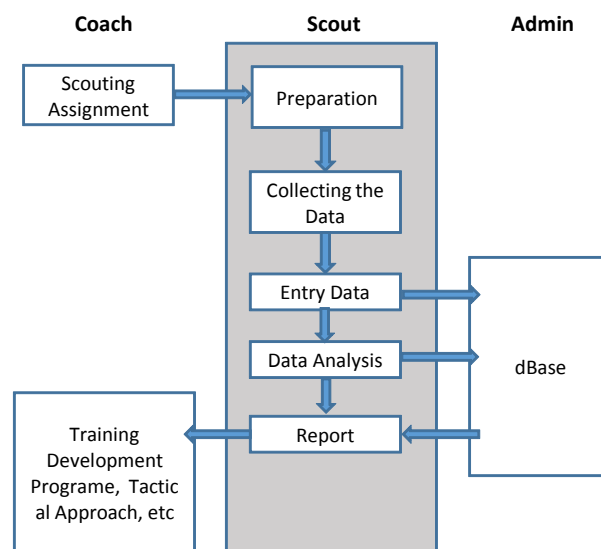

Figure 1. Data modeling to catagorized athlete's stroke
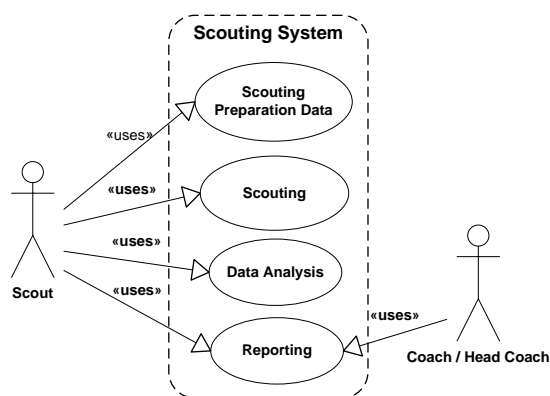

Figure 2. Scouting business process

Figure 3. Use case diagram
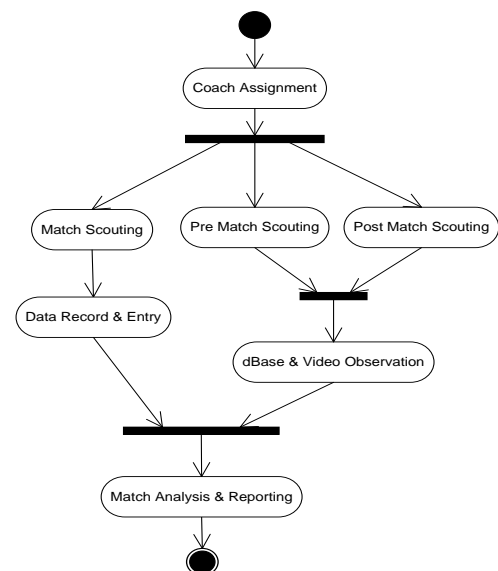
Figure 4. Activity diagram

In this research, match preparation conducted as Fig.2 scouting business process. The assignment of a scout will be conducted by coach or head coach, usually. Then the data will be collected for each match or game, based on each player with an opponent or event. Recently this activity using manual process will be replaced with the desktop based software. In the future, by using this software, it support

with an interface to player dBase that managed by an admin. Use Case Diagram of developed software was shown in Fig. 3 For this software the major user was scout itself, due to simplify the manual activities recently. By considering that most of the scouting activity was outdoor and mobile, the desktop based programmed was preferred. That is the reason of using Visual Basic, besides of their simple and easy to use characteristics. Fig.4 is shown activity diagram for this scouting software, which consists of 3 major activities that can be accomplished by the scout.

4. Results

In this research, a post-match analysis based on video files was used. Data collection to analyze in this research was taken from a video file of Final Asia Man Badminton Team Championship commence in April 2018, held in Malaysia. The chosen match was between Indonesian Jonathan Christie against Shi Yuqi from People Republic of China. Then the data were simulated using Tanagra ver. 1.4 to validate the result. The results are stroke pattern from Indonesian player and tactical approach analysis that implement during the game.

After observing the video matches, in the first set, there were 37 pieces item set omitted. Consist of 7 types of technical stroke. Starting from service, netting, drop shot, lob, smash, backhand and drive (return). Table 1, show the data statistic from SET 1, for instance, JC need 6.03 stroke for each point. Data also collected for the following sets, and then the analysis results were compared each other (Table 2 and Table 3) in order to figure out the game pattern for overall athlete's performance and tactical approach.

Table 1. Scouting card resume.

SCOUTING CARD RESUME							
Date	: unknown						
Event	: Final Badminton Asia Team Championship 2018						
Venue	: Stadium Abd. Halim, Alor Setar, Malaysia						
Scouting	: Jonathan Christie (INA)						
Opponent	: Shi Yuqi (CHN)						
Game	: 1st Men Single / Set-1						
	Serve	**Netting**	**Drop Shot**	**Lob**	**Smash**	**Backhand**	**Drive Return**	**Total**
Total	17	61	46	43	18	20	18	223
Avg	0,46	1,65	1,24	1,16	0,49	0,54	0,49	6,03
Min	0	0	0	0	0	0	0	1
Max	1	5	7	7	3	4	2	22
Tactical	8%	27%	21%	19%	8%	9%	8%	100%

Figure 6 shows the Resume Game Pattern in Set 1. Netting (27%), Drop Shot (21%) and Lob (19%) are a dominant technique used by JC. Based on this statistical number, the technical approach used by JC is very prudent to start the match. Spare the energy for the long endurance game, on the match, with seeking an opportunity to "kill". The parameters used during the analysis, minimum support 0.25, in apriori frequent item set parameter.
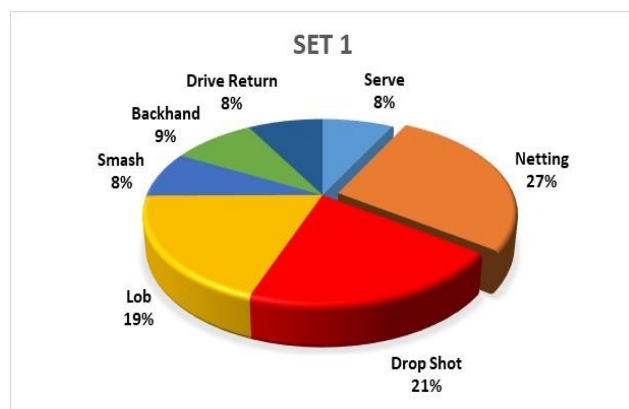

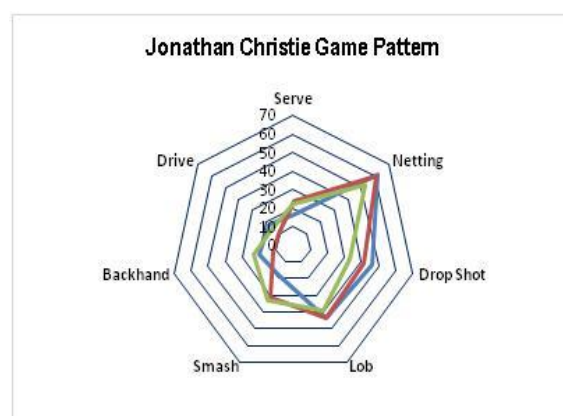**Figure 5.** JC's Resume game pattern in set 1

Figure 6. JC's game pattern comparation

Figure 7 clearly provides a graphical display of the JC game pattern for 3 sets. It can be seen that the game area of set 1 has a shift in set 2, as well as set 3. After defeating in set 1, JC made an improvement and different approach so he won set 2 and set 3. The blue color indicates set 1 pattern, less of smash and backhand. The red for set 2, and grey for set 3, they were indicates JC play more aggressively. A badminton athlete does not only need physical and technical abilities to compete at the highest level of competition. But also requires the strong character in order to influence the way of opponent plays, while not being affected by the way the opponent want to play. Besides that, it also requires creativity to give the opponent a shocking effect. That was what JC had shown in that match, by changing his playing rhythm and pattern toward an unwilling situation of his opponent (in set 2 and 3).

Table 2. Resume card for Set 1, 2,3.

SET	Serve	Netting	Drop Shot	Lob	Smash	Backhand	Drive	Total
SET 1	17	61	46	43	18	20	18	223
SET 2	24	60	41	43	31	12	11	222
SET 3	22	52	33	39	33	23	15	217

Table 2 indicates, the total stroke per set were almost the same, it is around 220 strokes (with a deviation of less than 1 %). This shows the consistency of the two players in this match.

Table 3. Changes of game pattern

SET	Serve	Netting	Drop Shot	Lob	Smash	Backhand	Drive	Total
SET 1	8%	27%	21%	19%	8%	9%	8%	100%
SET 2	11%	27%	18%	19%	14%	5%	5%	100%
SET 3	10%	24%	15%	18%	15%	11%	7%	100%

Table 3 illustrates data changes in the game pattern (tactical approach) carried out by JC. JC was consistently doing accurate netting, diverting variations from a drop shot by increasing the smash intensity and playing effectively to win points by point. In addition, JC opponent was seen trying to make a change by pressing JC's backhand side especially in the 3rd set. But this was anticipated by the JC (or JC's coach team). JC also has better stamina and mental (fighting spirit) preparation, became the key factor to victory for JC.

5. Conclusion

FP-Growth has the ability to provide an overview of the game pattern or tactical approach of badminton athlete(s) in a game. Based on calculation and analysis using FP-Growth, the JC game pattern is very monotonous in set 1. JC relies more on Lob, Netting and Drop Shoot to defeat the opponent. Therefore JC made changes on set 2 and 3 (see Table 2 and 3). Not only relied on a drop shot, began to smash and play more aggressively, especially in set 3 (see in Fig 6). Knowing the opponent game pattern and tactical planning is very important to win the match.

At the high level of competition, non-technical preparation can be a determining factor. This is especially when meeting opponents who have equal balanced technical abilities. JC defeated in 1st set, is a result of many factors. These factors include mental, self-motivation, detail knowledge, field adaptation, understanding of the opponent's behavior/character, and the ability to improvise against the opponent's condition. Sport science has become a necessity to develop in order to increase JC (Indonesian athletes) competitiveness.

6. References

[1]   T. O. Bompa, *Periodization: theory and methodology of training. 4th ed.* 1999.
[2]   Merriam-Webster Unabridged, "Merriam-Webster Dictionary," *USA*, p. https://www.merriam-

webster.com, 2019.

[3]   M. Phomsoupha and G. Laffaye, "The Science of Badminton: Game Characteristics, Anthropometry, Physiology, Visual Fitness and Biomechanics," *Sport. Med.*, vol. 45, no. 4, pp. 473–495, 2015.

[4]   R. M. Anggraeni, "Perbandingan Algoritma Apriori dan Algoritma FP-Growth Untuk Perekomendasi Pada Transaksi Peminjaman Buku di Perpustakaan Universitas Dian Nuswantoro," *Tek. Inform.*, pp. 1–5, 2014.

[5]   A. Ikhwan, D. Nofriansyah, and Sriani, "Penerapan Data Mining dengan Algoritma Fp-Growth untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus Kampus STMIK Triguna Dharma)," *Saintikom*, vol. 14, no. 3, pp. 211–226, 2015.

[6]   M. S. U. Pang-Ning Tan, U. of M. Michael Steinbach, and U. of M. Vipin Kumar, "Association Analysis: Basic Concepts and Algorithms," *Introd. to Data Min.*, pp. 238–414, 2006.

[7]   B. Santosa, "B . Santosa , ' Data Mining : Teknik Pemanfaatan Data untuk Keperluan Bisnis , Graha Ilmu ', Edisi Pertama , Cetakan Pertama Graha Ilmu , Candi Gebang Permai Blok R / 6 Yogyakarta," *Data Min. Tek. Pemanfaat. Data untuk Keperluan Bisnis*, vol. Edisi Pert, no. Graha Ilmu, pp. 2012–2013, 2017.

[8]   A. Eko and M. R. Ratih, "Penggunaan Algoritma FP_Growth Untuk Menemukan Aturan Asosiasi Pada Data Transaksi Penjualan Obat di Apotek (Studi Kasus : APOTEK UAD)," *J. Sarj. Tek. Inform.*, vol. 2, pp. 130–139, 2014.

[9]   J. Han, M. Kamber, and J. Pei, *Data Transformation by Normalization*. 2011.

[10]   D. Samuel, "Penerapan Stuktur FP-Tree dan Algoritma FP-Growth dalam Optimasi Penentuan Frequent Itemset," p. 6, 2008.