Historical Relative Performance Index over Interconnectedness of Badminton Athletes

Deni Khanafiah[†], Rolan Mauludy[††]. Hokky Situngkir[†††]
Dept. Computational Sociology
Bandung Fe Institute
&
Research Division - Surya Research International

[†] *mail: dk@compsoc.bandungfe.net,* [††]*mail: rmd@ compsoc.bandungfe.net,* [†††]*mail: hs@compsoc.bandungfe.net*

June 16th 2007

Abstract

The paper proposes the Historical Relative Performance Index in order to quantitatively extract information in the scores hit in the sets of head-to-head game in badminton tournaments. The index is treated as the weights of the directed networks built between competing athletes. The paper also proposes the way to build the fully connected network based on the empirically found network in order to have relative index between athletes that have never nor will be met in series of games. Some further directions as well as implementation to small amount of data is described for advanced analysis.

Keywords: badminton tournament, athlete's performance, network of athletes.

1. Introduction

Probably, there is no other aspect of human modern life that is more competitive than the world of sport. However, sport grows with certain rules and with the invigorating of the fairness in its development from one tournament to another. However, the tournament system in our age has grown better and better, all improvement was made for the sake of open possibility that only the best athlete would become the champion. From the large varieties of sport, some of them are played as a collective game and it needs a very complex teamwork and not solely by individual skill in order to win a game. However, other large varieties of the game of sports are played individually or if it incorporates teamwork, it did in doubles or very limited number of players. In this kind of game, the head-to-head meets among opposite players is very important not to mention the importance to see the performance of an athlete from her historical meetings with other players.

Sports tournament, like badminton, tennis, squash, *et cetera*, records the meetings of players in a large number of tournaments. However, there are still not many analytical approach being proposed to extract any useful information that might exist in them. This is the motivation of the paper. We realize that the head-to-head game results are emerged from the complexities of player's condition at the exact moment of the game. This complexity comes from a lot of sources, psychological, social, even sometimes political. Simply speaking, to win a game in this kind of sport, the aspect of health fitness, intelligence, emotions, strategy, and techniques must be in the excellent state.

This paper reports an endeavor to extract the information in the historical chart of badminton tournament by proposing the Historical Relative Performance Index calculated from the scores gained by any players in sets of games in tournaments. This is followed by building the network of athlete based upon the meeting in various tournaments. In order to complete the network of index, we apply a model for relatively completing the unconnected athletes to build a fully connected network of athlete. Several further directions are also included as well as implementation to small amount of data of badminton scores in two seasons of Olympic Games.

2. Historical Relative Performance Index

The model is proposed by constructing a full-connected network between players in the dataset. The full-connected network is comprised by vertices that are representing the corresponding players and the edges between players that are representing the relative performance index, reflecting the aggregate results of any historical games between them. For instance, a tournament will yield a matrix of the head to head game results on each player. Of course not all players will have chance to meet with all other players in a single tournament, and it is possible that in a single tournament two players meet twice (or more) in the series of the games. A game might result two or three sets of games, and the scores each player gains are the parameter denoting the result of the game: who wins or loses.

Imagine we have some historical tournaments with their respective results for each set of games. For example, we have two competing teams, *A* and *B*, and each team would be represented by $\{A_1, A_2, A_3\}$ and $\{B_1, B_2, B_3\}$ respectively. A game i could be comprised of a number of sets (τ_i), and the relative strength of player x to player y can be calculated by the fraction of scores gained by player x from the total scores hit throughout the game. However, in one tournament, player x and player y might be met more than once. Thus, in a single tournament, we can build an adjacency matrix of players (K) which elements are the aggregate fractions of the scores gained by players,

$$k_{xy} = \frac{1}{N}\sum_{i}^{N}\left[\frac{1}{\tau^{(j)}}\sum_{j}^{\tau^{(j)}}\frac{k_x^{(j)}}{k_x^{(j)}+k_y^{(j)}}\right],\ k_{xy}\in K \qquad (1)$$

where k_x denotes the total score of player-*x,* N denotes how many times the two players meets in the single tournament. The value of k_{xy} would be in [0,1]. As the $k_{xy} \to 1$, the game is relatively more easier for player x. It obvious that player x won most sets of the game when $k_{xy} > 0.5$ and vice versa.

Furthermore, from T series of historical tournaments, we have the relative strength index between player and player y. Thus, we now have, the Historical Athlete's Relative Index between player x and player y over T tournaments that is calculated as,

$$\alpha_{xy} = \sum_{t=1}^{T}\xi^{(t)}k_{xy}^{(t)} \qquad (2)$$

and

$$\sum_{t=1}^{T}\xi^{(t)} = 1 \qquad (3)$$

where $\xi^{(t)}$ is the weight factor of each tournament we put into account which is event-dependent in the horizon of historical games[1]. Since, the Historical Athlete's Relative Index becomes the elements of an adjacency matrix ($\alpha_{xy} \in \mathrm{A}$) in the athlete's network, any new tournaments, with some possibilities of the newcomers in the event the new tournaments may update the adjacency matrix A' as,

$$\alpha'_{xy} = \frac{\alpha_{xy} + \xi' k_{xy}'}{1+\xi'} \qquad (4)$$

This updating rule from any new tournament in the history of the corresponding sports gives us possibility to measure the relative strength of an athlete over time in our athlete's network. Here lies the interesting part of the dynamic Historical Athlete's Relative Index.

[1] Obviously not all sport tournaments can be considered to be the same. For instance, the meeting between player x and player y in an Olympic round should be considered more important in the index relative to the one in a regional based tournament. The more important one game, the greater the value of $\xi^{(t)}$.

3. The Network of Games

From the previous series of calculations, we have the adjacency matrix of Athlete's Historical Relative Performance Index representing the relative qualities that shows how an athlete performed relative to one another in the historical chart. We might be able to draw a network of the games from this matrix and have a visual description of the empirical results of all the games not solely by the numbers of the winnings but by the scores gained in each game: the vertices denote the player and the edges denote the relative performance. From here, it is clear as longer the time of our reference, the more possible that a player does not have connection to other players. Mathematically speaking, the more $\alpha_{xy} = 0$ would be seen in our global matrix. An example is shown in figure 1.

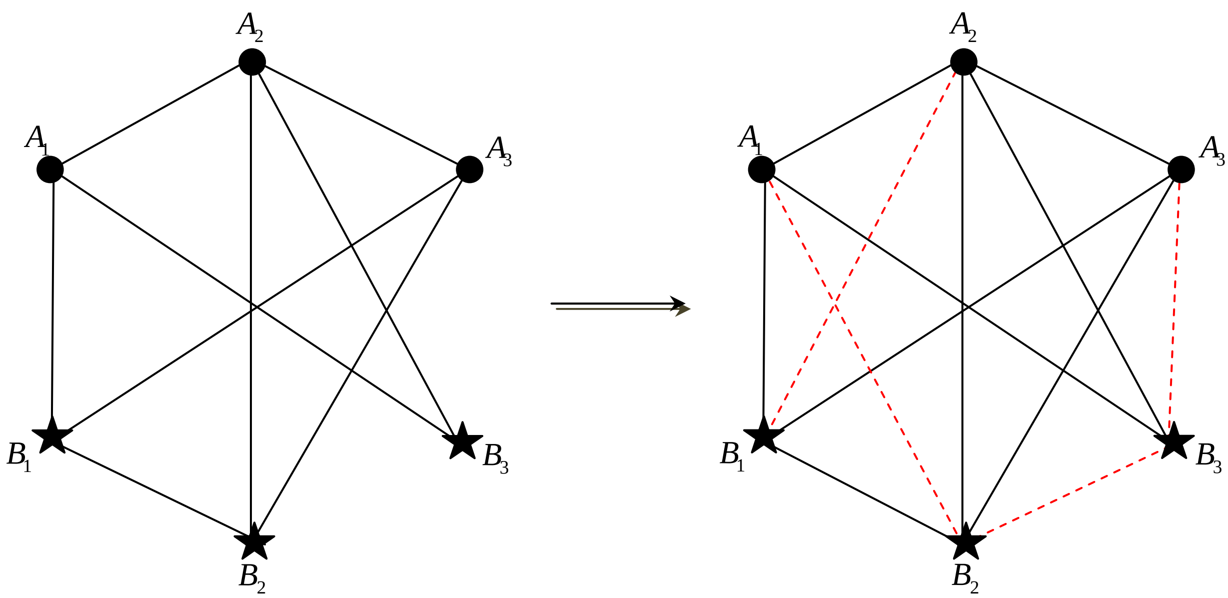

Figure 1
The transformation from the yielded historical relative performance network to the hypothetical relative performance network

It is tempting to have the hypothetical relative performance network that is defined as the fully connected network by filling the empty edges by certain algorithmic calculations based on the existing historical relative performance network. We do this by adopting equation of the chain chemical enzymatic reactions as theoretically proposed by biochemists Michaelis-Menten (*cf.* Mathews, *et. al.*, 2000). From figure 1, we see that the red-dashed lines are the hypothetical edges made up from our calculations.

The question is how we can 'predict' the more likely values of the red-dashed lines. To answer this, we refer to the previous section that the historical network was comprised by the sets of matrices of which elements are the aggregate performance index between two players *x* and *y* (α_{xy}). We must realize, at long enough periods most players should play in many series of tournaments and events, and thus most players have ever met each other. If player *x* wins most games over *y* in several games in a single or more tournaments, and player *y* wins most games over *z*, intuitively, we could say that *x* would wins over *z* if both eventually meet in a game. Hence, *x* plays relatively better than *z* from this simple historical chain of games. By using this intuitive thinking, we build an analogous model with the one used in chains of chemical reactions widely understood and used in biochemistry. From this point of view, we could say that "the winning of *x* over *y* would be a 'catalyst' on the possibility of the winning probability of *x* over *z*" and vice versa there would be some cases that the probable result of a game plays as 'inhibitor' over other games. When we met a practical case of $\alpha_{xy} > \alpha_{yz}$, it is more likely that there would be a game that reveals $\alpha_{xz} > 0.5$. This has been discussed a lot in the enzymatic reaction in biochemistry. Default enzymatic reaction can be written as:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} P \qquad (5)$$

that show how one substrate react with certain enzyme (be it catalyst or inhibitor) to produce particular product. Based upon this model we construct our model with some modification of this default model, i.e. with assumption that reaction in all steps is reversible. If we have a game of *X*, *Y*, and *Z*, that can be written as reversible first-order reaction-like as

$$X \underset{k_{yx}}{\overset{k_{xy}}{\rightleftharpoons}} Y \underset{k_{zy}}{\overset{k_{yz}}{\rightleftharpoons}} Z \qquad (6)$$

then we could have the relations of formation of $[Y]$ as

$$[Y] = k_{xy}[X] + k_{zy}[Z] \qquad (7)$$

and since we concern with the formation of $[X]$, we can get

$$[X] = k_{yx}[Y] \qquad (8)$$

Thus, substitution equation (7) to equation (8) we can have

$$\frac{[X]}{k_{yx}} = k_{xy}[X] + k_{yx}[Z] \qquad (9)$$

then at equilibrium condition, the formation of $[X]$ represented by $[Z]$ is calculated as:

$$[X] = \frac{k_{zy}}{\left(\frac{1}{k_{yx}} - k_{xy}\right)}[Z] \qquad (10)$$

The ratio of rate constant then represent into single constant k_{xz} as:

$$k_{xz} = \frac{k_{zy}}{\left(\frac{1}{k_{yx}} - k_{xy}\right)} = \frac{k_{zy}}{\left(1 - k_{xy}k_{yx}\right)} \qquad (11)$$

Thus, in our cases of which $k_{xy} \approx 1 - k_{yx}$, we have,

$$k_{xz} \approx \frac{(1-k_{yz})(1-k_{xy})}{1-k_{xy}(1-k_{xy})} \tag{12}$$

By the same calculation, for a longer chain reaction:

$$X_1 \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} X_2 \ldots \underset{k_{-N}}{\overset{k_N}{\rightleftarrows}} X_N \tag{13}$$

We can get:

$$[X_1] = \frac{\sum_{i=2}^{N} k_i}{1 - \sum_{i=1}^{N} k_i k_{-i}} [X_N] \tag{14}$$

In chemistry, the symbol of [...] represents the concentration of the particular chemicals, and rate constant (k) represents the velocity of reaction of particular products at a given certain concentration of reactants. In our analysis the value of [*X*] can be denoted as unity. However, in further analysis, as we realize that there are so many "X-Factor" when we talk about sport concerning emotional, physical, and many other (qualitative and quantitative) wholesomeness of an athlete as she performs in a game, the symbol might also be incorporated as the percentage potentials that she exhibits as the game was on the stage. The similar model has also been incorporated to analyze the competing firms in duopoly games concerning the levels of advertising exhibited by firms (*cf.* Situngkir, 2006).

The above model is performing only for single element in each step. In chemical reaction, it is also known as first order reaction. However, chemical reaction involves a lot of chemicals and can be representing as *n*-order reaction, for instance,

$$\alpha_1 A_1 + \alpha_2 A_2 + \ldots \alpha_n A_n \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \beta_1 A'_2 + \beta_2 A'_2 + \ldots \beta_m A'_m, \tag{15}$$

with rate of formation can be calculated as

$$v = k_1 [A_1]^{\alpha_1} [A_2]^{\alpha_2} \ldots [A_n]^{\alpha_n} \tag{16}$$

Even though it is seemingly to be more sophisticated, the *n*-order chemical reaction model can give us some possibility to develop further model above that is involving not only a single entity but many. Furthermore, large varieties impact of each entity to the global system are also possibly different, for example the similar analytical model for the sports played by groups of people, or even further, the collective action of people in particular system.

In advance, by finding the shortest paths that includes the maximum total of the Historical Relative Performance Index, we can calculate the more likely values of the hypothetical one. For instance, to find the non-existing α_{xz} , we must find the shortest path with maximum sums of the Historical Relative Performance Index from all possible *k* numbers of path in the existing network, *N*.

$$\alpha_{xz}^{hypothetical} = \max\left(\alpha_{x1}^{k} + \alpha_{x2}^{k} + ... + \alpha_{nz}^{k}\right),\ \forall k \in N \tag{17}$$

The shortest path can be found by using the Dijkstra's algorithm (*cf.* Cormen, *et. al.,* 2000) over the adjacency matrix (A).For instance, we use the Dijkstra's algorithm to find the shortest route between two players that is actually never met in our historical tournaments. From this perspective, we could hypothetically measure the relative strength between player *a* and *b*, even though they never met before as we calculate the relative strength from the game between *a* vs *c*, and *c* vs *b*.

Thus, by using these transformative calculations, we could fill most of the zeros in the Historical Relative Performance matrix and we easily measure the strength of certain teams as a tournament is coming along. From the eventually yielded fully connected graph, we could find two qualitative measurements, i.e.: the likeliness relative qualities of the players in our game. This can be measured respect to time (the degrading or upgrading trends of each player in one team or simply by the comparative of relative performance among players in the opposing teams. Nonetheless, the proposed algorithm might also be used to becoming an alternative index reflecting an athlete's performance over a series of competitions or tournaments. This interesting issue is left for interested readers.

4. An exemplification from Indonesian Badminton Teams

Badminton is a kind of sport that is very interesting in the perspective of Historical Relative Performance Index. Indonesia has been well known for her world class players in many international stages of badminton tournaments, *e.g.:* All England Tournaments, Sudirman Cup Tournament, including universal sport seasons like Asian Games and the Olympic Games. The single male or woman of a badminton game somehow shows the quality of the players and the emergent results are more likely to reflect the quality of players at the exact moment of the game. From the head-to-head data of Summer Olympic Games 2000 and 2004[2], we apply the algorithmic steps as described in the previous section and see the results of the game in a view of historical chart of complex network. This is shown in figure 2. The famous and tight competitions between Indonesian and Chinese teams are shown as the highlighted vertices of the network.

[2] http://en.wikipedia.org/wiki/Badminton_at_the_2000_Summer_Olympics/ and http://en.wikipedia.org/wiki/Badminton_at_the_2004_Summer_Olympics_-_Men's_singles/

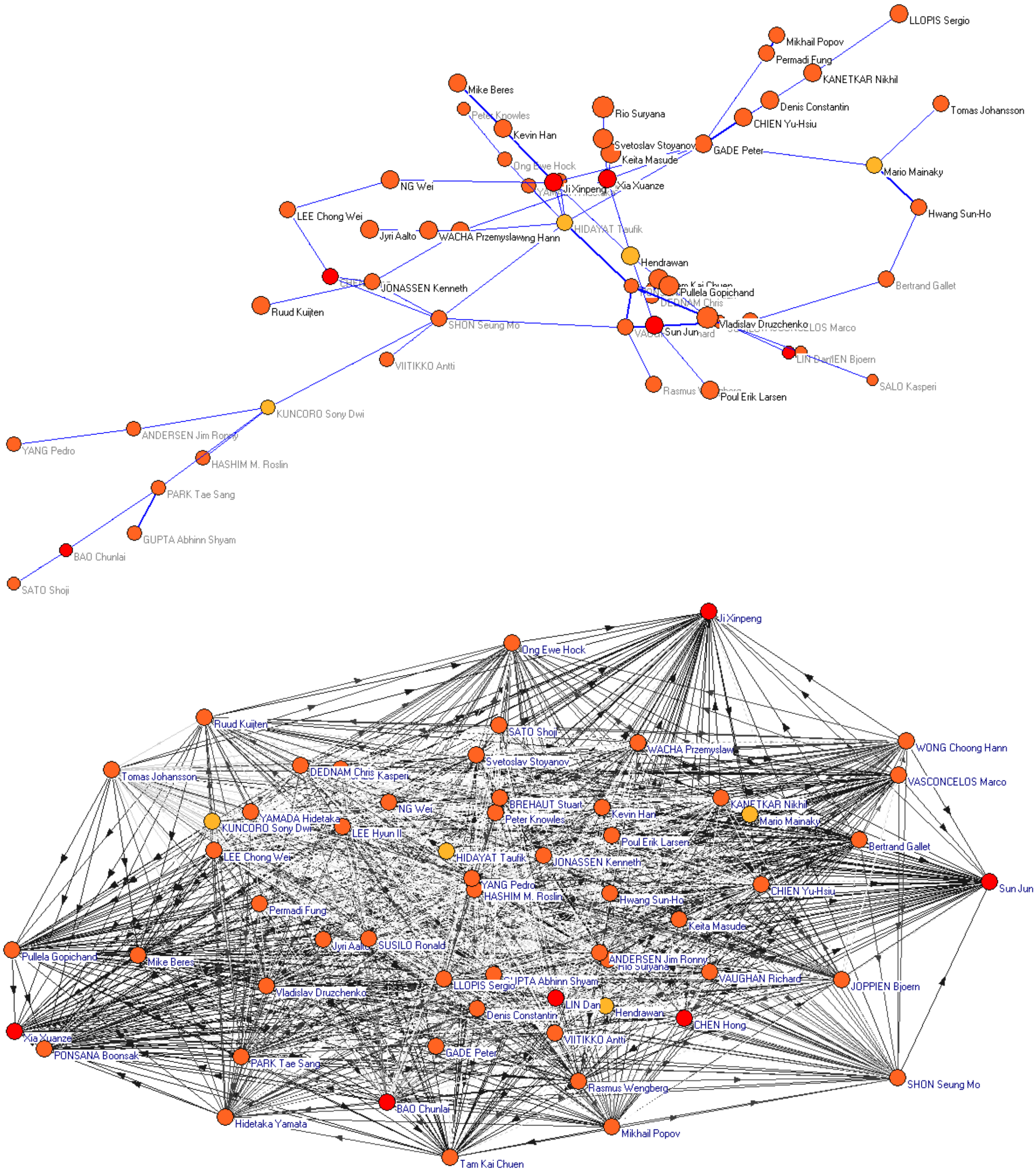

Figure 2

The bidirected network of badminton male athletes meeting in Summer Olympics 2000 and 2004 (*above*);
and the one with the extrapolations by using Historical Relative Performance Index. The dashed line means those with negative value of $\hat{\alpha}_{xy}$.

The network shown in figure 2 is constructed by using the directed graph that is obtained from the relations,

$$\hat{\alpha}_{xy} = \left(\alpha_{xy} - \alpha_{yx}\right) \qquad (18)$$

in order to reduce the coefficient obtained from the equation (18). In the network, the directing arrow of $\hat{\alpha}_{xy}$ shows the position of the edges in the matrix: $x \rightarrow y$ and the negative values show that at this case player y is more likely to win the series of games in the tournaments. Figure 2 shows interestingly the information that we might have by incorporating the Historical Relative Performance Index as we can observe the relative qualities of the games presented by the players.

In a glance view, we could clearly see that the Indonesian and Chinese teams were tightly competed in the two Olympic Seasons. Players in both teams are connected with most of the players participating in the Olympic Games reflecting that both of them are more likely to advance any final rounds regarding to the adopted tournament system. Interestingly, from the 58 payers in the two tournaments both national teams are well connected with players from other teams by means of the more likely to win edge representations. There are a lot of usefulness that we can gain from this understanding as the numbers of the data (*i.e.:* numbers of players and the tournaments) employed in the model for the sake of evaluative purposes before or after any tournaments.

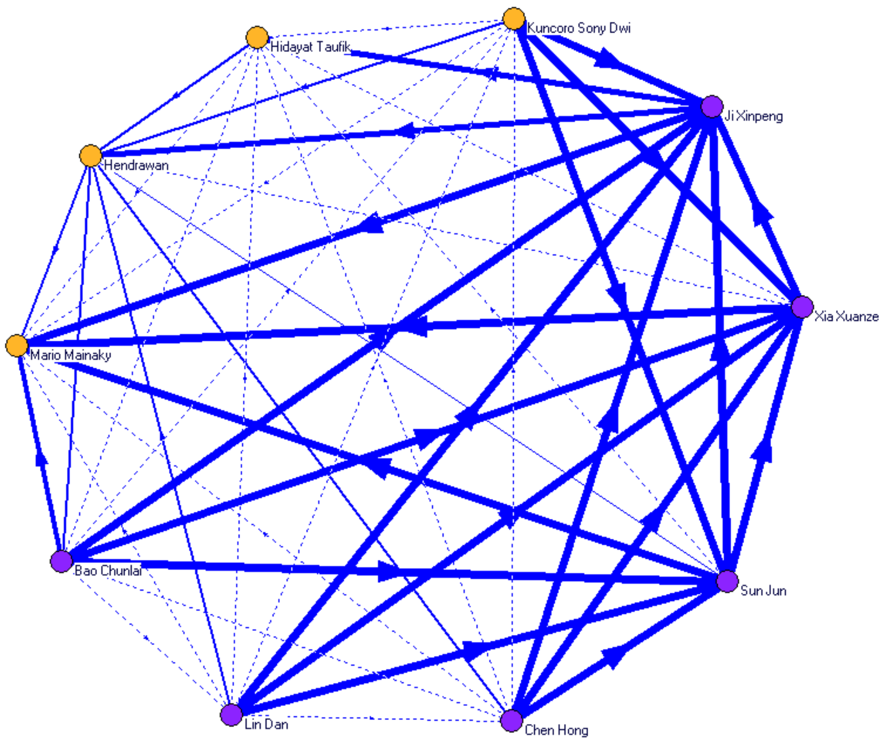

Figure 3
The relative strength between Indonesia national badminton team with badminton team from China from their performance in the Olympic Games 2000 and 2004.

A more figurative visualization of the exemplification between badminton team of Indonesia relative to their counterpart from China is shown in figure 3. The thickness of the arrows reflecting the bigger relative value while the arrows shows the first player to be notated in the model. For example, the relative strength between Jin Peng (CHN) and Hendrawan (INA) is shown by solid arrows from Ji Xinpeng to Hendrawan. This shows that Ji Xinpeng is relatively stronger than Hendrawan as evaluated from the performance in the two Olympic Games. This strong differences is however smaller by looking at the arrow from Hendrawan to Sun Jun (CHN), as the solid line is thinner. The dashed line is interpreted in the opposite way, the arrow comes to the vertices of one player reflects that the player is relatively stronger. As the analyzed data grows, the better observation might be able to be delivered in comparing two or more badminton teams.

5. Ending notes and Further Directions

We propose the use of the Historical Relative Performance Index of athletes in badminton tournaments as a way to extract valuable information from the historical data of games among athletes. The idea is to construct an athlete network comprised by the graph of which vertices representing athletes participating in certain tournaments and edges connecting them to those with the corresponding head to head games. The model is constructed along with the capability for the updating of the data purposes as new tournaments are coming. The inspiration from the catalysts-inhibitors in biochemical enzymatic reactions are also incorporated in the model in order to build the fully connected network among athletes.

From the built model, we can observe a lot of things, for instance the relative strength between an athlete with another that can be adapted to give dynamic evaluative information of the relative performance throughout time and the comparative analysis of one player with another from other teams or country. The implementation is adapted to the Olympic Games results of badminton tournaments which the head to head games and complex individual capabilities are playing very vital role for the teams in the sustainability in particular tournaments.

However, further directions of this approach can also be directed to the statistical properties of the athlete's network as it has been pioneered in a lot of works *e.g.*: the network of scientific collaboration network (Newman, 2000), film actor network (Amaral, *et. al.*, 2000), and a lot more growing researches on complex social network. This opens a big door for further directions originated in this report.

Acknowledgement

Authors and all researchers of Surya Research International regarding to this report thank Tan Djoe Hok for moral support and discussions. The paper is dedicated to the improvement of the world of badminton in Indonesia.

Works Cited:

Amaral, L. A. N., Scala, A., Barthélémy, M., & Stanley, H. E. (2000). "Classes of Small-world Networks". *PNAS* 97 (21): 11149-52.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2001). *Introduction to Algorithms* 2nd Edition. MIT Press and McGraw-Hill.

Mathews, C. K., van Holde, K. E., & Ahern, K. G. (2000). *Biochemistry* 3rd Edition. Addison-Wesley.

Newman, M. E. J. (2001). "The Structure of Scientific Collaboration Networks". *PNAS* 98 (2): 404-9.

Situngkir, H. (2006). "Advertising in Duopoly Game". *BFI Working Paper Series* WPF2006.