Application of T-Pattern Detection and Analysis in Sports Research

Gudberg K. Jonsson1, M. Teresa Anguera2, Pedro Sánchez-Algarra2, Conceição Olivera3, Jorge Campanico3,*, Marta Castañer4, Carlota Torrents4, Mária Dinušová4, Javier Chaverri4, Oleguer Camerino4 and Magnus S. Magnusson5

1Human Behavior Laboratory, University of Iceland, Iceland & Department of Psychology, University of Aberdeen, UK
2Department of Methodology of the Behavioural Sciences, University of Barcelona, Spain
3Department of Sport Sciences, Exercise and Health, University of Trás-os-Montes e Alto Douro, Portugal
4Human Motricity Laboratory - INEFC, University of Lleida, Spain
5Human Behavior Laboratory, University of Iceland, Iceland

Abstract: The following paper introduces a new approach to the analysis of sports and motor skill performance. The approach, known as T-pattern detection, is explained and preliminary data analyses from soccer, boxing, basketball, swimming and motor skill analysis are presented. The data show that specific temporal patterns can be identified within sports performances. The temporal patterns can relate to performance of specific actions (e.g. passes) or movement patterns. Further analysis showed a significant correlation between coaches’ ratings of performance with the number of temporal patterns exhibited by athletes.

Keywords: Pattern detection, pattern analysis, T-patterns, Theme, soccer, boxing, basketball, swimming and motor skills.

INTRODUCTION

The main drive of performance analysis is based on the premise that coaches and teachers cannot observe and recall all the critical, discrete incidents from a sport and motor skills performance [1]. They should therefore benefit from the provision of additional information that describes a performance at a level of detail that they cannot recall. It is assumed that provision of quantitative information on performance will impact on the coaching process through an enhancement of the coaches’ interpretation of performance. Additional information can also enhance performance through improved quality of performer feedback, provided that the feedback is in appropriate forms [2].

In the pursuit of generating quantitative information on performance analysts have traditionally used frequency of event occurrence as their index of performance e.g. the analyst has recorded how many passes have been made from particular playing zones or how many times possession has been lost. In essence the analyst has been answering the question “how many times did ‘x’ occur?” However frequency of event occurrence has been shown to be an inadequate index of performance that cannot differentiate between effective performances [3]. If one accepts the argument that sport performance consists of a complex series of interrelationships between a vast array of performance variables then simple frequency data can only ever provide a relatively superficial view of performance.

If performance analysis is to continue to advance understanding of sport and motor skills performance then it must continue to explore better methods of collecting and analyzing data. The purpose of this paper is to introduce and explain a new data analysis method that has the potential to make a significant contribution to analyses of sports performance. Data from preliminary studies of sport and motor skill performance are also presented to show the potential outcome from the analysis process.

T-PATTERN DETECTION AND ANALYSIS

The new analysis approach being presented is based on a process known as T-pattern detection which allows for detection of the temporal and sequential structure of a data set. The method has been developed, outside of sport, on the assumption that complex streams of human behaviour, such as sport performance, have a temporal/sequential structure that cannot be fully detected through unaided observation or with the help of standard statistical and behaviour analysis methods. Given that observational records of human behaviour, including sport performance analysis, have both a temporal and sequential structure an analysis tool that can describe this structure will enhance understanding of the behaviour (s) being studied. A generic observational software package called Theme has been specifically developed to operationalise T-pattern detection as an analysis process [4-6].

A schematic representation of a T-pattern is shown in Fig. (1). If one assumes that the letters in line 1 correspond
to specific performance events (e.g. pass, tackle & shot in soccer) that appear on the line in proportion to the time of their occurrence then line 1 is a visual representation of the temporal structure of a sports performance.

Within the upper line there are four events \((a,b,c,d)\) that occur in a regular temporal pattern however the pattern has been masked by the surrounding, more random, occurrence of the events \(w\) and \(k\). If a performance analyst or coach were simply visually inspecting the data string it is unlikely that the pattern would have been detected. Equally no form of frequency count, lag-sequential analysis or time-series analysis would have identified the event pattern. However a T-pattern analysis would have identified the pattern because of it’s consistent temporal structure. The T-pattern detection algorithms allow an analyst to separate out randomly occurring events from temporal patterns even when the random events occur in between elements of the pattern.

The most essential aspects of T-pattern detection will now be presented to give some insight into the theoretical base of the process. A complete explanation of the theoretical roots of the pattern-detection algorithms together with an overview of the wider use of the process has been presented elsewhere [4, 5].

### T-Pattern Definition and Detection

A T-pattern is essentially a combination of events where the events occur in the same order with the consecutive time distances between consecutive pattern components remaining relatively invariant with respect to an expectation assuming, as a null hypothesis, that each component is independently and randomly distributed over time. As stated by Magnusson ‘that is, if A is an earlier and B a later component of the same recurring T-pattern then after an occurrence of A at \(t\), there is an interval \([t+d_1, t+d_2]\) \((d_2 \geq d_1 \geq d_0)\) that tends to contain at least one occurrence of B more often than would be expected by chance’ [5, p. 94] The temporal relationship between A and B is defined as a critical interval and this concept lies at the centre of the pattern detection algorithms.

The pattern detection algorithms can analyze both ordinal and temporal data however, for the algorithms to generate the most meaningful analyses the raw data must be time coded i.e. an event must be coded according to time of occurrence as well as event type. The coding of many event-types and corresponding times results in the type of data set shown in Fig. (2). This figure displays a behaviour record from the second half of a club soccer match and consists of 250 series of occurrence times (one for each coded event type) ordered according to their first occurrence time.

Within the analysis process, in addition to identifying patterns by their temporal characteristics, it is also assumed that hierarchical relationships exist between patterns. The pattern in Fig. (1) shows two smaller patterns \((a-b \& c-d)\) combining to produce the larger pattern \((a-b-c-d)\). It is as-

![Fig. (1). Schematic representation of a T-pattern viewed within a normal data string and as it appears in isolation.](image)

![Fig. (2). A time series behavior record from the second half of a soccer match from the European Champions League 1997. The match was coded from a digitized video recording of approximately 45 minutes duration (time is in seconds).](image)
assumed that a hierarchically organized temporal pattern type will have a self-similar structure in that a complex T-pattern is essentially a pattern of patterns of patterns etc.

The assumption of a hierarchical structure to patterns is necessary to make pattern detection a workable analytical process. Even in moderate data sets the number of potential T-patterns is very high. When, for example, the potential number of event codings is 100, the number of potential event patterns involving up to 10 event codes is many orders of magnitude greater than $10^{11}$ if all possible time windows are also considered. Even for supercomputers it becomes an impossible task to search for each possible temporal pattern separately. To deal with this problem a bottom-up, level-by-level detection strategy, based on the assumed hierarchical structure of T-patterns, is used to detect complex patterns as patterns of simpler temporal patterns.

The bottom-up detection algorithm involves two main stages. The first deals with a search for the critical interval relationships, as defined earlier, in the identification and construction of new patterns. The second stage deals with completeness competition between all the detected patterns. In this stage those patterns which are less complete versions of one or more alternate patterns are deleted. For example pattern $Q_x$ is considered less complete than pattern $Q_y$ if $Q_x$ and $Q_y$ occur equally often and all events in $Q_x$ also occur in $Q_y$. In this case $Q_x$ is eliminated from the analysis since it provides no useful additional information. The completeness competition ensures that only the most complete patterns survive and constitute the result of the detection process.

Advantages and Characteristics of T-Pattern Analysis

Negating the Influence of Intervening Behaviours

The number and type of behaviours that may occur between the components of a T-pattern can vary greatly between occurrences of the same pattern. For this reason, methods that depend only on the order of events disregarding the temporal distances between them have great difficulty detecting such patterns. The same is true for methods that do not consider multi-ordinal or hierarchical relations between pattern components. For example, in a soccer match an attack down the right side of the pitch by team A may be temporally related to the same team conceding a corner a short time later. In between making the attack and conceding a corner team A may have lost possession in many number of ways. However the attack and the concession of the corner were causally related in that the attack has unbalanced team A’s defence resulting in their opponent making an effective counterattack upon regaining possession.

Current data collection approaches combined with statistical analysis methods currently used in major statistical packages as well as in specialized behaviour research, software such as The Observer [7, 8] or GSEQ [9, 10], would not identify the pattern described above. The variability in loss of possession in the middle of the pattern would mean that the pattern is only detectable in terms of its temporal characteristics. (Analysis of the characteristics of standard observational methods which make them inadequate for T-pattern detection, has been described elsewhere [5]).

A Cyclical Patterns

While T-patterns can occur cyclically it is not an essential characteristic. The distances between occurrences of patterns, not events, may just as well be irregular i.e. a pattern of play may occur three times within the first twenty minutes of a soccer match and then not reoccur until the last 5 minutes. The within-pattern relationships between events will have remained relatively invariant but overall pattern occurrence will have been variable.

Causality

Causality is also an issue that may or may not relate to any identified pattern. If one considers a simple tennis related example this issue is easily highlighted. If a tennis player bounces the ball prior to each serve and when the ball is served wide to the opponents backhand the return is played down the line then this is likely to constitute a temporal pattern. However within this pattern there are both non-causal and potentially causal relationships. The act of bouncing of the ball will not have been causally related to the direction of the serve. However the direction of the serve may have been causally related to the direction of the return i.e. the opponent can only play shots down the line when the ball is served wide. One cannot assume that just because a pattern exists the elements within it are causally related. This issue will be referred to again when preliminary data is discussed.

Prediction

The final feature of T-pattern analysis relevant to sport is the capacity to generate forward and backward conditional probabilities for any pattern that is identified i.e. the probability that when A occurs at time $t$ B will occur within a critical interval defined as $[t+d_1, t+d_2]$. Once patterns have been identified it is relatively simple statistically to identify probabilities thereby giving the coach and/or performer the potential opportunity to predict an opponent’s performance.

In essence T-pattern detection provides an integrated system that allows for increased depth of analysis and description of sports performance.

EXEMPLAR DATA

Methods

Preliminary investigations in soccer, boxing, basketball, swimming and motor skills have reinforced the author’s belief that T-pattern identification has great potential as an effective research tool in performance analysis. The soccer and basketball research has utilized multiple game analyses whereas the work in boxing, motor skills and swimming involved the case-study of a single performances.

Soccer

Thirteen soccer matches, five club and nine international matches were coded using a combination of the soccer match analysis system developed at Liverpool John Moores University and Theme. Coding included data on pitch position, player and match events. Pitch position was classified according to the pitch division shown in Fig. (3a). The primary event categories for data collection were: pass; tackle; header; run; dribble; clearance; shot; cross; set-play; lost
control; foul. Additional qualifying statements could be tagged to each event category. All data was analyzed using the Theme software package.

Boxing

A world championship boxing bout was coded and analyzed using Theme. The events coded were fighter, ring position, movement type and type of punch. Ring position was coded according to the ring divisions shown in Fig. (3b). The fight was stopped in the fourth round due to a knockout.

Swimming

The empirical data were obtained by using the coding of thirty breaststroke cycles, swum at maximal speed, of a national champion swimmer video recorded underwater in several moments. The video images were captured from front and side-view by classical underwater criteria and converted the stroke cycle to a digital format for analysis.

Basketball

The current study is a part of a broader research project concerning various team sports in both national and international competitions. Here ten games from the Spanish ACB League and the Final Four of the Euro-league played during the 2005-2006 season have been selected for analysis. The procedure was in line with ethics of APA and approved by the university departments involved. The project did not involve any experiments or manipulation of subjects. The results are based on data obtained from recordings from public TV. Table 1 shows in capital letters the team observed in each game. The objective here was to conduct an observational analysis of basketball, using the SOBL-1 observation instrument [11], in order to determine the influences on the game result considering both the dynamics of play, or interaction contexts [12], and the spatial distribution of players in terms of laterality.

RESULTS AND DISCUSSION

The studies show that a high number of complex temporal patterns exist in all of the data types analyzed, some occurring in cyclical fashion.

Soccer

A typical event pattern from the soccer analysis is shown in Fig. (4a). This figure displays a detected T-pattern that has four event types and occurred three times during the first half of a European Championship qualifying match (1998). The three boxes in this figure show the same observation period. The upper-left box shows the hierarchical construction of the pattern, while the upper-right box displays the time point of each event in the pattern and their pattern connection. The bottom box shows the hierarchical structure of the pattern

Motor Skills

The twelve participants were recruited from among the total of 120 first-year Sport and Physical Activity Sciences undergraduates from Lleida University (INEFC-UdL). The age range of the participants was 19-21. Although they had considerable experience in physical activities they had no prior experience concerning dance classes. Participants were video recorded during 7 lessons of Body Movement, each lasting fifty minutes and taking place over a period of two months. The main activities were based on motor improvisations, mime-dance and contact-improvisation. The instructor randomly used the three varieties of instructions (descriptive, metaphoric and kinesics) to explain the different tasks. In each task, participants had to explore, according to their own motor ideas, the instructions given. All sessions were video recorded after an adaptation phase to the camera, the aim here being to avoid any reactive effect prior to beginning the study period.

Table 1. List of Basketball Games Observed: The Team Observed in Each Game is Shown in Capital Letters

<table>
<thead>
<tr>
<th>Result</th>
<th>Venue</th>
<th>Teams</th>
</tr>
</thead>
<tbody>
<tr>
<td>Won</td>
<td></td>
<td></td>
</tr>
<tr>
<td>87-79</td>
<td>Home</td>
<td>MANRESA-Menorca</td>
</tr>
<tr>
<td>73-90</td>
<td>Away</td>
<td>Fuenlabrada-MANRESA</td>
</tr>
<tr>
<td>86-79</td>
<td>Home</td>
<td>MANRESA-Estudiantes</td>
</tr>
<tr>
<td>89-81</td>
<td>Home</td>
<td>MANRESA-Unicaja</td>
</tr>
<tr>
<td>78-81</td>
<td>Away</td>
<td>Leon-MANRESA</td>
</tr>
</tbody>
</table>
expressed in relation to the temporal structure (only complete patterns are shown in this box).

The pattern describes (Fig. 4b) how player A moves the ball towards the opponents goal by receiving the ball in, and then passing it out of, pitch zones 8, 11 and then 14 consecutively. Player A then completes the sequence by passing it on to player B who receives it in zone 15. The pattern describes an attacking movement through the middle of the pitch which opponents would clearly wish to prevent. Traditional frequency analysis of passing would have identified the ball reception and subsequent pass from each zone as discrete events but would not have linked the consecutive actions in the four zones. The movement from zone 11 to 14 also occurred on another five occasions during the first half (Fig. 4, upper right box) further suggesting that player A was working effectively through the central channel of the pitch. This integrated form of analysis can only enhance the information flowing to the coach.

In addition to immediate analysis of individual matches the data were also used to look at two additional issues relating to structure within team performance. The first issue considered was the comparative level of temporal structure within club and international soccer matches. In a simple data manipulation three randomly selected club and three international matches were compared in terms of the mean number of patterns and the mean number of pattern occurrences identified in each match type.

The data (Fig. 5) show that international football has a more defined temporal structure than club football. This finding may be due to the presence of higher technical abilities in international footballers which help create a more structured game or, alternatively, contextual differences between club and international football e.g. club football is played at a higher pace throughout mitigating against the development of structure within the game. Whatever the reason the clear difference in temporal structure between club and international soccer merits further investigation.

The second additional issue that was investigated related to the potential interrelationship between performance rating by coaches and the degree of structure in team performance. Three experienced soccer coaches observed one club and two international matches and were asked to rate the performance of every player (on both teams) on a simple ten point Likert type scale. For each coach the player ratings for a specific team were averaged to give a team performance rating. Team performance ratings were then correlated (Pearson product-moment correlation) against the number of patterns exhibited by each team. The data show that the coaches’ ratings of team performance were significantly correlated to the number of patterns identified for each team (r=0.81, p<0.05).

The link between performance rating and pattern participation suggests that coaches were recognizing, albeit at a potentially subconscious level, the structure within a team’s play. However the traditional rationale for performance analysis is that coaches cannot observe and remember discrete events within critical event sequences [1]. Yet, in this sample, the fact that coach performance ratings were correlated with pattern participation suggests that coaches were perceiving information about the interrelationships between events. This finding also warrants further investigation since it relates to such a fundamental foundation in the performance analysis literature.

Boxing

Boxing, in contrast to soccer, is an individual sport in which the nature of the competition dictates that the actions of the two performers will be intrinsically inter-related e.g. fighter A does $x$ therefore fighter B does $y$. The boxing data clearly show the potential for T-pattern analysis to generate meaningful information in this type of sport. In the bout that was analyzed a traditional frequency based analysis of the fight found no meaningful differences between fighters punching profiles e.g. there was no significant difference between fighters in terms of punches that hit ($\chi^2=2.324$, p>0.05) and punches that missed ($\chi^2=2.321$, p>0.05). However the T-pattern analysis identified meaningful event sequences, such as that shown in Fig. (6), in which a specific event sequence, that occurred four times during the first

Fig. (4a, b). A temporal pattern relating to attacking movement of the ball through the centre of the pitch. (3a). Data output from Theme analysis software showing temporal and hierarchical representation of a T-pattern. (3b) Schematic representation of the same data 1 Player A receives the ball in Zone 8, passes the ball to a team mate and runs forward. 2 Player A receives the ball in Zone 11, passes the ball to a team mate and runs forward. 3 Player A receives the ball in Zone 14, passes the ball to a team mate in Zone 15.
round, resulted in fighter A getting a clear right-hand punch to his opponents head. Given that there were no significant differences in punching frequencies between fighters the data suggest that the key to successful performance lies in the creation of sequences of events rather than attaining overall thresholds for event frequencies.

In addition to the patterns involving punches thrown the data analysis also showed a highly consistent ring movement on the part of one fighter. Fig. (7) and (8) show the temporal and spatial distribution of ring movement for fighter A immediately after the start of each round. The data show that in all four rounds the first twelve ring movements were spatially and temporally consistent. Clearly if this information can be made available to coaches then it becomes possible to develop strategies to disrupt such preferred movement patterns.

Overall whilst the boxing data represents a single case-study the identified consistencies in action and movement highlight the capacity of T-pattern analysis to provide more meaningful performance data than simple frequency analysis. If this data were fed into the coaching process it is likely that performance would be enhanced.

**Swimming**

The aim of the study was to introduce a data analysis method that examines temporal structure and interrelationships between events (movements) within breaststroke swimming actions. This analysis technique can identify consistent temporal patterns that exist within the flow of behaviour and thus provide a different view of the complex interrelationships between movements. To illustrate the results complete (Fig. 9) and incomplete (Fig. 10) patterns obtained...
Application of T-Pattern Detection and Analysis in Sports Research

The Open Sports Sciences Journal, 2010, Volume 3

at two different moments have been selected, each composed of three cycles. Complete patterns are those that integrate all seven phases of the breaststroke cycle and incomplete patterns are those that only present some of the phases.

Basketball

The analysis of patterns detected revealed that certain patterns were found to exist only in games won and others only in games lost. Others were found to occur across the two different game results. Figs. (11 and 12) are two examples of patterns detected, revealing repeated configuration of game events, not only occurring in the same order but also with significantly similar time intervals between each event occurrence. A total of 773 patterns were detected in the subset of games won (see example in Fig. 11) and 677 T-patterns in the subset of games lost (see example in Fig. 12).

Motor Skills

The study analyzed the diversity of motor skills related to three different kinds of instructions: descriptive, metaphoric and kinesic, with a special emphasis on the detection of time patterns. Using observational methodology and technology applied to movement the aim of the study was to adapt the Observational instrument of Motor Skills (OSMOS) [13] and create an instrument capable of analyzing the motor skill responses generated in lessons of Body Movement and Dance.

Analysis of the T-patterns obtained from tasks introduced using a kinesic instruction shows that the most frequent pattern detected is the relationship between responses that reproduce the instruction (Me), followed by motor responses that change one of the characteristics (Mt), usually time (t) or body posture or gestures (c), where these are the most susceptible to change. However, when we focus on the motor skills criteria related to the type of instructions, it can be seen that very rich T-patterns are detected when using a kinesic instruction. It seems that this type of instruction offers more ideas and possibilities to participants. One relevant T-pattern (see Fig. 13) shows that although participants vary some characteristics of the tendency (Mt), they still try different skills. They also seek to explore different motor skill combinations related to the instruction (m) by using stability and manipulation skills (m,Es-m,MI-m,Mc). Participants also produced responses that do not agree with the proposal (a).
Fig. (9). Description of all details of a cycle pattern. Each branch includes simultaneous occurrence of leg, trunk arms and head codes (01) first leg propulsion action phase; (02) first leg propulsion action phase; (03) second leg propulsion action phase; (04) first arm propulsion action; (05) second arm propulsion action; (06) first phase of arm recovery; (07) second phase of arm recovery; (08) end of cycle. Among the two swimming cycles we can observe small variations introduced by the swimmer.

Fig. (10). Pattern A includes only the branches of (01) first leg propulsion action phase; (02) second arm propulsion action; (03) first phase of arm recovery; (04) end of cycle. Pattern B includes only the branches of (01) first arm propulsion action; (02) second phase of arm recovery; (03) end of cycle.

Fig. (11). Example of a detected T-pattern found in all games won, minimum occurrences 3 (p.005). The pattern describes: (1) player arriving in laterality left in advanced position in respect to the opponent, situated in the middle line (iz,am); (2) the player with the ball returns to the centre (ce,mm); (3) player retrieves in centre and the opponent advances (ce,ra); (4) player retrieves in laterality right and the opponent advances team sets up the position on the court (de,ra); (5) player advances in laterality right and the opponent retrieves (de,mm); (6) player retrieves to the centre (ce,oo); (7) player return to laterality right (de,oo); (8) player passes the ball to laterality left in advanced position (iz,ar); (9) player passes the ball in laterality left retrieving (iz,mm); and (10) player retrieves in laterality left in respect to last position (iz,ra).
Fig. (12). Example of a detected T-pattern, the first of the 677 found in games lost, minimum occurrences 3 (p.005). The pattern describes:
(1) player advances to the centre (ce,oo); (2) the player pass the ball to laterality right in advanced position, otherwise, if this don’t turns into a fastbreak, this is a common transition offense, it could be a fastbreak because defense is not established) (de,ar); (3) player positioned in laterality right retrieving to the middle position in respect to the other team (de,mm); (4) player retrieves to the centre, the offensive team sets up the play and the team on defense guards each player (ce,ra); (5) player retrieves in laterality left in respect to last position (iz,ra); (6) the player pass the ball to laterality right in advanced position (de,ar); (7) player passes the ball in laterality left retrieving (iz,mm); (8) player retrieves to the centre (ce,ra); (9) player passes the ball to laterality left in advanced position (iz,ar); and (10) player advances to the centre (ce).

Fig. (13). T-pattern related to kinesics instruction. This relevant T-pattern (p<0.005) shows for the first criteria of instruction (marked by a square) that the participants vary some characteristics of the initial instruction (Mi). They also try to explore other motor skill combinations related to the instruction (m) and also appears very frequently the exact instruction (Me). An example of motor skills criteria we can observe (marked by a circle) in this T-pattern can be when a participant makes a static rebalance (Es) then runs (Ls) and jumps (Ed), then stops (a). It continues in contact with another partner (Mi) to do turns (Ea) and also combining impact manipulation (Mi) and conduction manipulation (Mc) skills with the partner.
Generally, when a kinesic instruction is used, the participants’ initial response corresponds to a tendency instruction or an exact instruction. They then try to vary the skills used (Mi), and also to vary completely the instruction (m). In this case, the most frequently-used motor skills are those related with manipulation. An example of the motor sequence for this T-pattern is described in Fig. (13).

CONCLUSIONS

The preliminary data shown in this paper highlights the potential for T-pattern analysis to make a significant contribution to sport performance and motor skills analysis. Current analytical methods that focus on simple frequency analysis cannot identify the temporal patterns within a sports and motor skills performance. Consequently without this form of analysis meaningful information is not being made available to the coach/teacher. If this information is not available then it possible that performance is not being optimized.

The data also point towards the need to investigate the potential link between temporal structure in sport and motor skill performance and the understanding of performance being generated by coach/teachers observations. The data suggest that whilst coaches/teachers may not be able to accurately recall discrete events they do perceive inter-relationships between events. This analysis approach can assist in generating a greater understanding of coach/teachers knowledge construction.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the Spanish government project Avances tecnológicos y metodológicos en la automatización de estudios observacionales en deporte (Dirección General de Investigación, Ministerio de Ciencia e Innovación) [Grant number PSI2008-01179]. We gratefully acknowledge the support of the Catalan government project Grup de recerca i innovació en dissenys (GRID). Tecnologia i aplicació multimedia i digital als dissenys observacionals [Grant number 2009 SGR829].

REFERENCES