Extracting Tennis Statistics from Wireless Sensing Environments

Adel Shaeib
Interoperable Systems Group

Kenneth Conroy
CLARITY: Centre for Sensor Web Technologies

Mark Roantree
Interoperable Systems Group

Dublin City University, Glasnevin, Dublin 9, Ireland
{ashaeib,kconroy,mark.roantree}@computing.dcu.ie

ABSTRACT

Creating statistics from sporting events is now widespread with most efforts to automate this process using various sensor devices. The problem with many of these statistical applications is that they require proprietary applications to process the sensed data and there is rarely an option to express a wide range of query types. Instead, applications tend to contain built-in queries with predefined outputs. In the research presented in this paper, data from a wireless network is converted to a structured and highly interoperable format to facilitate user queries by expressing high level queries in a standard database language and automatically generating the results required by coaches.

Keywords

XML, Query, Ubisense, Sensor

1. INTRODUCTION AND MOTIVATION

The TennisSense project is a collaboration between Tennis Ireland [5], who provide the user requirements for the system; CLARITY, a Sensor Web research facility that provides sensor devices; and the Interoperable Systems Group of DCU. In this research paper, we focus on the requirement to detect serves, break down the match into a series of game boundaries, and provide a standard query interface. One of the attempts used to automatically generate the information required by coaches is the deployment of Ubisense [6], a wireless sensor network that tracks objects or persons wearing Ubitags. This acts as an indoor localisation system that can track and capture an objects three dimensional location in space. In our case, the Ubitags were carried by both tennis players involved in a series of matches.

This paper presents our methodology for automated processing of wireless sensor data so that it can be interrogated using a standard query language. The major benefit of our research is to provide queryable statistics to coaches immediately after matches so that they can evaluate the sensed data and modify the behaviour of their athletes for subsequent matches. The contribution of our work is the automated enrichment of low level sensor data together with a methodology for extracting the required results. Our approach is novel in that our automated process for enriching raw data is applied to a semantic transformation process for user queries. This employs the concept of tennis court zones and match states to express the queries of sport scientists and coaches. A more detailed description of this system can be found in the technical report[4].

DMSN'09, August 24, 2009, Lyon, France

2. SENSOR ARCHITECTURE

The TennisSense architecture has three major component processors. Process 1 (P1) takes raw Ubisense data and converts to XML; P2 performs a semantic enrichment of the data using the semantics of the game of tennis; and P3 transforms high level queries so that they can execute on the enriched Ubisense data.

Ubisense generates data based on players' movements. When a player moves, Ubisense detects the Ubitag carried by the player and transmits a time stamp, Ubitag id, and Ubitag location which is player's (x,y,z) coordinates. The role of the Ubisense Processor is to provide a readable and queryable dataset from raw Ubisense data. This is performed in two passes: the first pass provides a structural enrichment and creates a single XML document for both players (for the entire match); the second pass enriches the file with zone information specific to the tennis court, and also detects when players change ends.

3. DATA PROCESSING AND ENRICHMENT

The Ubitags generate (x,y,z) coordinates for the current position of the device, and therefore the player, on the court. In order to interpret meaning for these values, the tennis court is mapped into 24 zones, each a range of (x,y) coordinates corresponding to a section of the court, as illustrated in figure 1. During the enrichment of Ubisense data, the court zone for every Ubisense instance is noted in an XML tag, as are special zones where applicable. Categorising the player location on court helps in the construction of generic queries and the detection of similar patterns of play. The court mapping is a necessary step to allow basic queries in XPath and XQuery following enrichment.

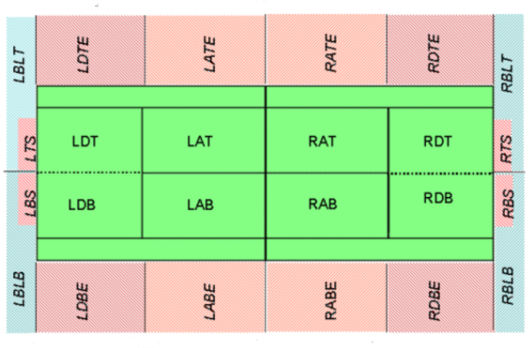

Figure 1: Court divided into 24 zones.

4. QUERY PROCESSING

The query requirements of tennis coaches can be expressed in XPath or XQuery. An example query in XQuery is shown below.

Example 1. How many data values (Ubisense) did player 1 record during the match?
let $a := doc("fileName.xml")/match/player[@id="1"]
return string($a/@entries)

5. EXPERIMENTAL DATA

Experiments ran on a desktop computer of Intel(R) Core(TM) 2 Duo, 2.66GHz, 3.25GB of RAM, Windows XP Professional SP3 operating system, and with Java virtual machine 1.6 and eXist database version 1.2.5 was implemented as XML files host.

The first step in the verification process was to compare the `detectServe` results with the actual serves in a match. We randomly picked a single best-of-5-set match containing a total of 87 serves. `detectServe` detected a total of 65 serves, representing 74.7% of the total. 16 of these detections should have been marked as two serves from the same zone (in quick succession). This represents a further 18.3% of the total number of serves with an incorrect enrichment tag. Two serves (2.2% of the total) were missed completely by the system, and there were four instances where a serve was incorrectly identified.

For performance reasons, the time taken by the system to process raw data files is measured, as is the time taken to structurally enrich the file into XML and the time taken to semantically enrich the file with zone and state information. All files gathered so far have length times between 18 minutes and 28 minutes, and sizes between 500KB and 1,400KB. The time taken to enrich the file in its entirety, from reading it in to full semantic enrichment even in the largest file, is under 900ms.

Two types of test were performed regarding high-level queries. The first was the blanket detect serve algorithm which consists of a mixture of XQuery and Java and returns a list of all serves detected during a match. The other type of query consisted of XQuery alone, such as the detection of all serves in game one. The detect serve algorithm is as a rule computed on-the-fly during semantic enrichment and a resulting `Serve.xml` is the result. It is on this file that we perform subsequent detect serve queries, hence results are returned extremely quickly (<20ms).

6. RELATED RESEARCH

Related work in the tennis domain has focused on extracting highlights of tennis matches by using video and audio information[1] [3]. The benefit of our approach is that the sensor network is less computationally expensive than the video network and we provide a queryable database. The research presented in [2] describes SMARTLab, an OSGi-based middleware platform for intelligent environments. The focus of SMARTLab is on the middleware platform, of which Ubisense only forms a small part. Our approach differs as we apply the semantics of a tennis match to enrich sensor data and provide a database for user queries. In [7], multiple sensor data, including Ubisense data, is used to detect behavioural patterns. However, they do not address the enrichment of data, nor the requirement for user queries.

7. CONCLUSIONS

In this paper, we described our efforts to use a wireless sensor network to provide tennis coaches with query interface to tennis match statistics. Unlike other approaches (using video), the Ubisense network is less expensive both computationally and in terms of cost, and even at this point, we can demonstrate a high level of accuracy when detecting key events in match data. Future work will focus on complete automation of annotating the game boundaries by customising the `detectNewGame` algorithm. We also concentrate on providing queries for velocity and acceleration of players. This will assist in the identification of further patterns of play such as baseline rallies or fast passages of play.

Acknowledgement

This work is partly supported by Science Foundation Ireland under grant 07/CE/I1147.

8. REFERENCES

[1] Huang Y-P. and Chiou C-L. and Sandnes F.E. An intelligent strategy for the automatic detection of highlights in tennis video recordings, in *Expert Systems with Applications* (ESWA), 2009.

[2] López-de-Ipiña D., Almeida A., Aguilera U., Larizgoitia I., Laiseca X., Orduña P., Barbier A., Vazquez J.I. Dynamic discovery and semantic reasoning for next generation intelligent environments, in *Intelligent Environments* (IET), 2008.

[3] Pingali G.S., Jean Y., and Carlbom I. Real Time Tracking for Enhanced Tennis Broadcasts, in *Computer Vision and Pattern Recognition* (CVPR), 1998.

[4] Technical Report. http://www.computing.dcu.ie/ isg/publications/ISG-09-02.pdf

[5] Tennis Ireland. http://www.tennisireland.ie

[6] Ubisense English Site. http://www.ubisense.net/

[7] Yang G-Z., Lo B., Wang L.(J), Rans M., Thiemjarus S., Ng J. From sensor networks to behaviour profiling: a homecare perspective of intelligent buildings, in *IEEE Seminar on Sensor Systems for Intelligent Buildings*, 2004.