Trustworthy Collaborative Investigation System for Match Fixing in Cricket using Spectator Voting Scheme

K.Ananthapadmanabha* and Dr.K.Udayakumar**
*Research Scholar, PP COMP.SCI.ENG.0251
Computer Science & Engineering, Rayalaseema University, Kurnool, India. e-mail: siridevikap@yahoo.co.in
**Principal & Professor, Dept of Computer Science & Engineering,
Adarsha Institute of Technology, Bangalore, India. e-mail: aitbprincipal@gmail.com

Abstract: Cricket has a history of more than hundred years. In these years it has seen some extraordinary controversies like Body line series of 1930's, Ball tampering rows and latest one is that of match fixing in cricket. According to Kevin Carpenter the author of "Match fixing –The biggest threat to sport in the 21st century", the greatest threat to sport globally is match fixing. Match fixing in cricket is a rare event since not all cricket matches played are fixed. To verify rare events like match fixing in cricket infrequent pattern mining techniques are used.
Collaborative filtering for match fixing in cricket is used to make recommendation, observation and analysis about players and others involved in match fixing based on other players, expert commentators and police views, observation and previous experience. In this paper we are considering a trustworthy collaborative investigation system for match fixing in cricket using spectator voting scheme.

Keywords: Match fixing, Collaborative Investigation, Voting scheme.

Introduction

Cricket is a game of glorious uncertainties. Match fixing in cricket is a rare event. Out of N cricket matches played a very small number of matches say p are fixed, where as q number of matches are not fixed. q number of matches represents normal legitimate transactions and p number of matches represents fraudulent transactions. The objective of our research is not in predicting matches which are not fixed, but we are highly interested in predicting matches which are fixed even though their number is very small and their occurrence is very in frequent.
If our model indicates that one percent of the cricket matches played are fixed and fraudulent. It also indicates that 99 percent of the cricket matches are not fixed and they are legitimate cricket matches. These legitimate cricket matches has an accuracy of 99 percent but fails to detect any of the matches which are fraudulent and fixed.
Analyzing and detecting a rare event like match fixing in cricket is like finding a needle in a big ocean. Due to this reason there is a need to put up large efforts in designing models for rare events like match fixing in cricket. Existing data mining algorithm for frequent pattern mining like Apriori algorithm and frequent pattern growth algorithm will work effectively and with high accuracy for normal legitimate cricket matches which are not fixed but will fail to address rare events like match fixing in cricket.

Literature Survey

In the research paper titled "Forest fire model proposal for match fixing in cricket based on criminal network analysis" the authors focuses on analysis of match fixing based on areas of mining on social networks, link mining tasks. key tasks and challenges faced in this area are highlighted. A Graph model is used to represent the network and its related features.
The research paper " KL Cluster Nearest Neighbor Outlier Prediction Algorithm for Match fixing in cricket" highlights the importance of outlier analysis in a cricket match with N participants from which K clusters are created out of which there are L number of Match fixers.
In the research paper titled " Algorithmic Design Notation for match fixing in cricket using outlier analysis", the authors focuses on need for analyzing cricket match for verifying match fixing and proposes an algorithm called MatchFixers().
In the paper titled " Role of Iceberg diagrams as a Data Visualization Tool verifying match fixing in cricket ", the authors highlights how Iceberg Diagrams are highly suitable for multidimensional data analysis for mapping information on a two dimensional computer screen.
In the paper " Match fixing network analysis to verify nearness among internal participants of a cricket match " the authors focuses on role of Internal participants of a cricket match in match fixing and how they share critical game plan information from dressing room with bookies which help bookies in their betting transactions.

Infrequent patterns

An infrequent pattern is an participant set or a rule whose support is less than minimum support threshold value. Analysts find them interesting for rare event like match fixing in cricket, since not all matches played are fixed.

Key challenges in mining infrequent patterns

This includes issues like

i) How to identify interesting, in frequent patterns like match fixing in cricket
ii) How to successfully discover match fixing in cricket in large data
iii) What is the role of negative patterns, negatively correlated patterns and negative association rule in discovering match fixing in cricket.
iv) How to apply concept hierarchies for identifying match fixing in cricket.

Negative patterns

They contain negative participants sets and negative association rules.

Negative participant set

Negative participant set V is an participant set that has the following features

V = PP U NP, where PP is a set of positive participants and NP is a set of negative participants with the condition NP≥1 and support (PP)≥ minimum support.

Negative association rule

It is an association with

Methodology

Trustworthy collaborative investigation system makes use of collaborative problem solving mechanism which includes a trustworthy collaborative board whose members include

1. **Cricket experts** who are expert commentators with relevant domain knowledge about the cricket and match fixing in cricket.
2. **Police officers** from cyber crime department who are responsible for investing fraudulent activities like match fixing in cricket. They are well assisted by computer professionals from cyber crime investigation department.
3. **Expert Auditors** are responsible for verifying any illegal financial transactions involving cricket players, umpires and bookies. This is because match fixing is done for some monetary benefits through illegal financial transactions. If any such evidence is found auditors need to bring to the notice of police officers for further investigation.
4. **Cricket fans and spectators** are the most important stake holders of cricket. Their opinion about occurrence of match fixing in a cricket match is very crucial to obtain their opinion a spectator voting scheme is employed to check whether the match is fixed or not.

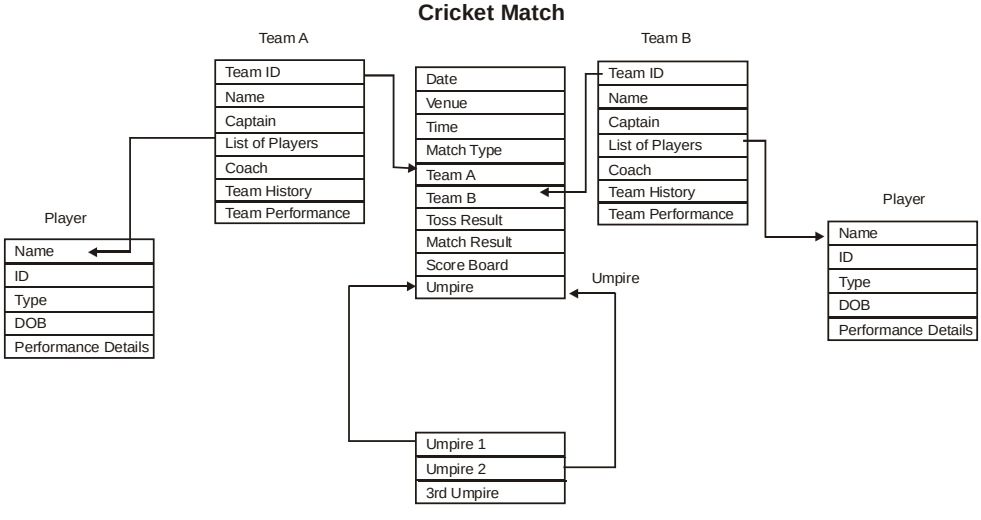


Figure 1: Relational Schema for a Cricket Match

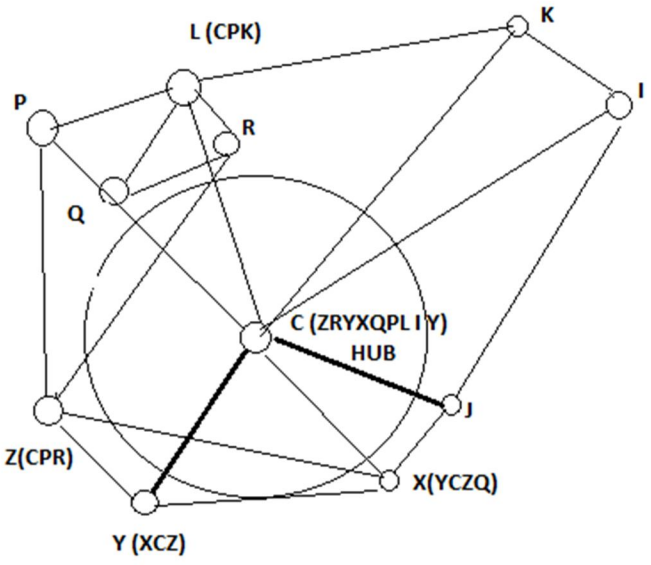


Figure 2: Criminal Network Diagram

Here a collaborative recommendation system is used to employ data mining techniques or statistical techniques to search for common similarities between different stake holders and come to a common conclusion. Two major challenges faced by collaborative filtering system are handling large volumes of data and basis for recommendations. This is required to gain public trust. If public in general follow the system but do not trust the system then they may not believe in the end result. The type of errors that a recommendation system can make includes false negative and false positive.

False negative are the facts that the system fails to identify and recommend although the general public would believe in it. False positive are the facts that the system recommend but which the general public does not like. Dimension reduction, association mining, clustering and Bayesian learning are some of the techniques used for collaborative recommendation systems. This system may consider information retrieval system which is a form of intelligent query answering system by analyzing the intent of the query and providing neighborhood information with respect to queries.

Expert commentators observe, analyze and identify any participant of a cricket match who is involved in match fixing and when there is any doubt of cricket match may be fixed. This information about match fixing will be brought to the notice of police officers for further investigation. Now the police officrs try to analyze criminal network formed between players, umpires, bookies and gangsters to extract criminal record information.

Expert auditors check for any illegal financial transactions by participants of the Cricket match and if there is any such activities it will be brought to the notice of police officers for further investigations.

The recommendations made by the investigation system should be trustworthy else the spectator will loose faith in the findings of the investigations. To check whether the investigations findings are trustworthy or not, spectator voting results and analysis are conducted. This helps in knowing the opinion of the spectators about occurrence of match fixing in cricket. For this a sample set of spectators is selected by looking into their history of audience ship i.e., how many matches they have been following over a period of time.

A online questionnaire is provided to selected spectators of N numbers and their opinion is collected in terms of issues like

1. Whether match is fixed or not: Yes/No
2. If fixed who the culprits are: list of players names provided.
3. The list of players who have underperformed during the match.
4. What event led to match fixing,
5. Any other related details if any.

The received information from the spectators is consolidated and corresponding weight is assigned per spectator. If high percentage of spectators agrees that the match is fixed, then this information is compared with investigation information from the collaborative investigation systems. If both of these matched then it is believed that there is enough evidence for match fixing.

If few spectators who have voted agree while others disagree that the match is fixed, then percentage of spectators who agree and percentage of spectators who disagree is calculated. This percentage of votes is compared with a threshold value. If the percentage of spectators agreeing is greater than threshold value then we compare information in investigation systems with this information.

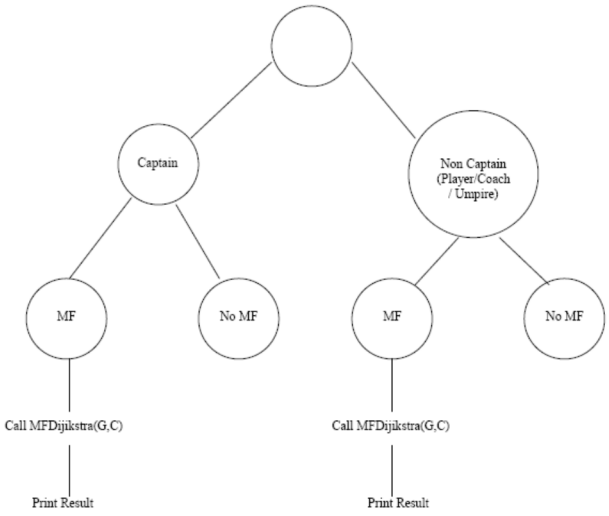


Figure 3: Decision tree for match fixing in cricket

Algorithm For Spectator Voting ()

Algorithm SVS( )
//Purpose: To get insight into match fixing from spectator prespective
//Input : Online Questionnaire provided to *n* selected spectators
//Output : Consolidated opinion of spectators about match fixing which is called SVS finding.

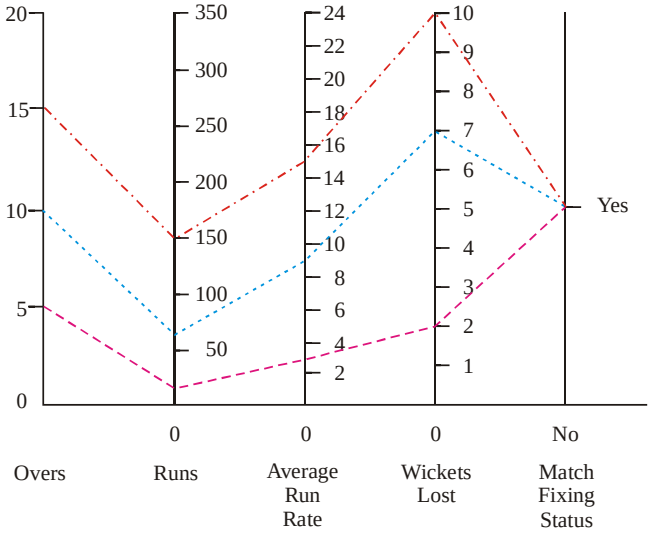
//Step 1: Spectator opinion extraction step.
Step 1: Collect the following information from *n* spectators using questionnaires

a. Whether they feel that match is fixed or not
   Yes Y
   No N
b. If match is fixed who are the suspects?
c. Who are the culprits?
d. What events led to match fixing?
e. Any other related details if any

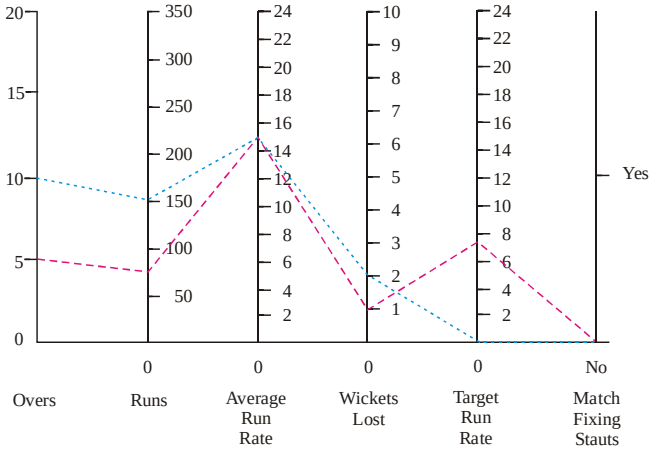
//Step 2: Spectator vote consolidation step.
Step 2: If high percentage of spectator agrees about match fixing
If Spectator match fixing status is Yes and total spectator vote exceeds threshold value.
    Send SVS finding to TRUSTWORTHY ( )
Else
If Spectator match fixing status is NO and Total No. of spectator voting is less than threshold value
Print “Match is not Fixed”
return SVS finding

Results & Discussion

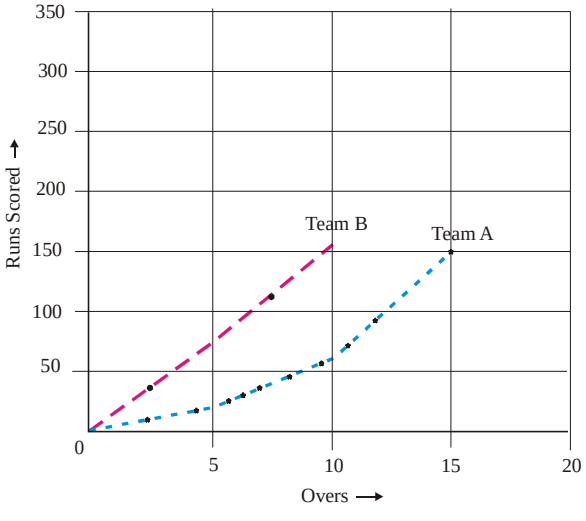
In a match between team A and team B, team A wins the toss and elects to bat first. As target setter team, team A scores 15 runs in the first five overs (overs 1-5) by loosing 2 wickets at an average run rate of 3 runs per over.
In the next five overs(overs 6-10)they score 45 runs at an average of 9 runs per over by loosing 5 wickets.
In the next five overs (overs 11-15) they score 90 runs at an average run rate of 18 runs per over by loosing 3 wickets. Team A sets a target score of 150 runs in 20 overs.
Graph 2 represents Iceberg diagram for team B as team chaser in their first five overs (overs 1 - 5) team B scores 75 runs at an average run rate of 15 runs per over by loosing 1 wicket. Their average run rate is above the target run rate required at this stage. In overs 6-10 team B scores 76 runs at an average run rate of 15 runs per over by loosing 1 wicket. Here also their run rate is above target run rate required . Team B wins the match in ten overs.



Graph 1: Iceberg Diagram for Team A (Target Setter) in a Match



Graph 2: Iceberg Diagram for Team B (Target Chaser) in a Match



Graph 3: Line Graph for Comparison between Team A and Team B Performance in a Match

From the line graph it is very clear that at all stages of the match, the chasing team, team B was performing much better than team A.

Conclusion

In this paper trustworthy collaborative filtering approach with audience voting scheme is applied which a very novel and innovative scheme for verifying match is fixing in cricket. Also issues like frequent pattern mining for identifying the match fixing in cricket. Since domain experts like ex cricketers and auditors are part of collaborative approach. It has enhanced the accuracy of results generated. Other issues like Simpsons paradox for avoiding generation of spurious patterns and reducing the time required for searching patterns is also considered. Involvement of spectators voting scheme helps in gaining insight of spectators view and opinion about match fixing in cricket.

Future Enhancements

To enhance trustworthiness of prediction , algorithm like Decision review system or DRS() can be included. If spectator voting scheme outcome do not match with domain experts results then DRS() can be invoked. Spectator voting scheme information may be collected or extracted from social medias like Twitter, Face book and Whats app.

Acknowledgements

Authors wish to thank the Management of Adarsha Institute of Technology, Bangalore and special thanks to Dr.P.V.Krupakara for their constant support.

References

[1] K. Udayakumar and K. Ananthapadmanabha, " KL Cluster Nearest Neighbor Outlier Prediction Algorithm for match fixing in cricket ", International Journal of Advances in Electronics and Computer science ISSN NO- 2393-2815, Special issue sept 2016

[2] Dr. K. Udayakumar and K. Ananthapadmanabha, "Forest Fire Model proposal for match fixing in cricket based on Criminal Network Analysis", International Journal of Engineering Research Volume no 5, Issue Special 4, pp-790-991, ISSN-2319-6890(Online), ISSN-2347-5013(Print), 20-may-2016.

[3] Dr. K. Udayakumar and K. Ananthapadmanabha, "Algorithmic Design Notation for Match fixing in cricket using Outlier analysis ", International Conference on Higher Education Conclave 2016 at Sheshadripuram First Grade College, Yelahanka, Bangalore held on 19th March 2016, ISBN No-978-81-8251-575-9.

[4] K. Ananthapadmanabha and Dr. K. Udayakumar , "Role of Iceberg Diagram as a Data Visualization tool for verifying match fixing in cricket ", National Conference on Recent trends in Computer Science held at Sheshadripuram First Grade College, Yelahanka, Bangalore held on 19th March 2016, ISBN No-978-81-8251-575-9.

[5] K. Ananthapadmanabha and Dr. K. Udayakumar , "Role of Iceberg Diagram as a Data Visualization tool for verifying match fixing in cricket ", National Conference on Recent trends in Computer Science and Engineering, MANTHANA-2017, SJCIT, Chikkabalpura, Karnataka on May-03-2017. ISBN NO-

[6] K. Ananthapadmanabha and Dr. K. Udayakumar, "Match Fixing Network analysis to verify Nearness among Internal participants of a cricket match ", IEEE International Conference on Electronics, K. Ananthapadmanabha and Dr. K. Udayakumar, "Match Fixing Network analysis to verify Nearness among Internal participants of a cricket match ", IEEE International Conference on Electronics Information Communication Technology", held on May 19-20 2017, SVEC, Bangalore, India,ISBN-978-5090-3704-9/$31.00@2017IEEE

[7] Anany Levitin, "Introduction to Design and Analysis of Algorithms", Pearson, 2nd edition, 2013.

[8] Alex Berson , Stephan J. Smith, "Data Ware Housing, Data Mining and OLAP", Tata McGraw Hill Edition publication

[9] R.V.Hauck et.al, "Using coplink to analyze criminal justice data", Computer, March, 2002,30-37.

[10] Pang-Ning Tan, Michael Steinback and Vipin Kumar, "Introduction to Data Mining", Pearson Education. 2015

[11] Arun K pujari, "Data Mining Techniques", 2nd edition, Universities press, 2009

[12] Jaiwei Han and Micheline Kamber, "Data Mining Concepts and techniques" 2nd Edition, Morgan Kaufman Publishers, 2006.