ISSN 2319-2720
Volume 3, No.1, April - June 2014
International Journal of Computing, Communications and Networking
Available Online at http://warse.org/pdfs/2014/ijccn02322014.pdf

A METHOD FOR MINING INFREQUENT CAUSAL ASSOCIATIONS WITH SWARM INTELLIGENCE OPTIMIZATION FOR FINDING ADVERSE DRUG REACTION

M.Poorani[1], P.Nithya[2] , B.Umamaheshwari [3],
[1]Research Scholar, Department of Computer Science, PSG college of arts and science, Coimbatore, Tamilnadu, India, poorani.muthu3@gmail.com
[2]Assistant Professor, Department of Computer Science, PSG college of arts and science , Coimbatore, Tamilnadu, India, nithi.selva@gmail.com
[3]Assistant Professor, Department of Computer Science, PSG college of arts and science , Coimbatore, Tamilnadu, India.

ABSTRACT*:* Due to the increasing growth of the population on now a day's analysis to medical drugs also plays most considerable process, finding the relationship among one drug to another drug are used to prevent unexpected outcomes of patients in efficient manner. Finding those relationships in efficient manner data mining plays most imperative role to mine relationship and their reactions in well organized manner. Though, mining these relationships is not easy task due to the complicatedness of confine causality amongst actions and the irregular natural world of the actions of concentration in this purpose. In order to overcome this problem in this paper proposed an efficient algorithm to mine their causal relationship in efficient manner. Specifically, we developed a new interestingness assess, restricted causal-leverage, based on hybrid fuzzy recognition-primed decision (HRPD) model. In this paper presents a novel work that initially creates a fuzzy membership function for that causal relationship among drug and their reactions of those selected patient records. In order to optimize fuzzy parameter by using swarm intelligence based Exponential Particle Swarm Optimization (EPSO) optimization framework .Each and every consider number of particles select a fuzzy membership function for identification of best causal relationship and their reactions of drug in efficient manner by using cue represented in fuzzy model . It improves the results of Mining Infrequent Causal Associations among drugs in well efficient manner because of using EPSO Algorithm for further analysis and examination present through drug safety professionals, progress the accuracy of system.

Keywords: Fuzzy recognition-primed decision (RPD) ,Adverse drug reactions, association rules, data mining algorithms, interestingness measure, recognition primed decision model. Exponential Particle Swarm Optimization (EPSO).

1. INTRODUCTION

Data mining is one of the major imperative mining methods to mine important data in well organized manner ,there are several number of data mining methods have been used in earlier work in order to analysis the results and mine most important data in sufficiently such as clustering, classification and feature selection algorithms .In order to conquer the problem of real world medical applications and discover exact information to analysis the results in efficient manner ,initially the data are preprocessed and extract important features and mine important data in well organized manner for every medical ,image processing application etc,..In this work we majorly focus on mining the drug and symptoms reaction relationship for signaling Adverse Drug Reactions (ADRs) that are related to health care applications in medical investigation process. Initially the ADR results are collected from anyone of the medical research hospitals and finds important potential ADR pairs are mined using mining algorithm and finally performance evaluation.

Significant motivation comprises intermission to workflow, not have of notion with the purpose of an adverse result has happen, and trepidation of proceedings [1]. Adverse reactions might go away unobserved in anticipation of lots of patients have been precious. For instance, it contains medicine greater than before heart assault hazard [2-3]. In contrast to analysis the results of health care records such as the following important services in this process are medical services, drug instruction, and hospital permission. These administrative of health care records provide detailed information about the patients and outcomes results of every patient in hospital records. They frequently cover considerable populations and are enthusiastically obtainable [4-5], they, when in confidence associated simultaneously, make available a precious reserve for in advance approaching addicted to definite patient care through potentially little point in time

lags [6] . For illustration, they present contingent verification intended for the discovery of approximately each and every one important ADRs. This might significantly improve existing ADR indication creation system, predominantly designed for the discovery of uncommon ADRs or individual's resultant in such conclusion as hospitalization, disability.

In order to perform the identification or detection ADR results the causal and uncausal relationship of the drug and symptoms of the patients are mined exactly using hybrid fuzzy recognition-primed decision (HFRPD) .In this Hybrid FRPD model the important fuzzy cue values for finding best ADRs pairs .In this EPSO optimization system the fuzzy cue parameters are optimized and it enhances the results of FRPD model .The optimized cue values of ADR pairs are applied to FRPD model to discover exact ADR pairs for further analysis of the incorrect drug used by anyone of the hospital . The major contribution of the work is discussed in section 3, before that in section 2 we study the existing methods used for mining of ADRs pairs and then discuss contribution of the work, then experimentation analysis of the work and finally conclude, remarks of the work.

2. BACKGROUND STUDY

Comprehensive information of ADRs plays a critical responsibility in preventing/avoiding difficult events [7]. For example, by means of ADR pattern similar to drug symptom, mechanized system be able to investigate health records to observe as well as identify adverse events [8-9]. Such pattern be able to be used toward establish at-risk patient groups [10], moreover preserve as well help practitioners fine-tune their diagnoses as well as prescriptions [11]. Hence, scientifically signaling as well as subsequently validate ADRs is of financial moreover social importance. This paper will concentrate on signaling ADRs, i.e., successfully generating ADR hypotheses. Data mining techniques have been apply used for health administration designed for health study data analysis [12-15].

The common performance of the PSO whilst Gaussian disseminated random noise is extra toward the fitness function as well as rotation of analysis search space with at random performed [16]. The investigational results show to facilitate the presentation of PSO remain successful with existence of noise as well as number of belongings, noise still help out the PSO maintain away from individual attentive in neighboring optima. The investigations PSO be estimate toward a noise-resistant alternative everywhere the main PSO round be adapted so that frequent assessment of like applicant explanation are all-inclusive toward develop estimation the specific fitness of this exacting clarification. The comparison well thought-out a number of numerical problems by means of additional noise, as well as unsubstantiated information of obstruction prevention by means of particular or extra robots. The noise Resistant EPSO show considerably better performance than the unique [17].

3. MINING INFREQUENT CAUSAL ASSOCIATIONS WITH HYBRID FUZZY RECOGNITION-PRIMED DECISION (HRPD) MODEL

Particle swarm optimization is one of the swarm intelligence based optimization algorithm to solve the problem of real worlds applications ,and optimization problems also solved simultaneously .Based on the assignments of the velocity values only each and every particles move towards from best location to another best location . Once best optimized location are found then automatically the location of each and particles position are updated using position update formula ,in this work we refer a EPSO (Exponential Particle Swarm Optimization) to improve searching speed of optimization algorithm to optimize the cue of fuzzy recognition model for examination of causal relationship of each and every patients . EPSO is the opposite technique designed for standardize the problems in optimization of fuzzy membership functions in discovery of relationships among drug and symptoms pairs in healthcare records. Before optimization of the fuzzy membership cue values for mining infrequent causal associations

Initially first collect the health care records in hospital records, for collected data are preprocessed by removing missing data ,irrelevant data and redundant data in records .Once the data are preprocessed then list of the number of drugs presented in preprocessed results $D(d_1, \ldots d_m)$ simultaneously list number of symptoms for those drugs $S(s_1, \ldots s_n)$.From this collected drugs and symptoms generate pairs ,each will represent corresponding patients results ,in this work we are only interested to analysis and mine the relationship among drug symptom pairs only. In order to mine casual relationship among adverse drug symptoms pairs, generate several number of association rules respectively for each pairs . In fuzzy representation model where $< X, Y >$ and $C < X, Y >$ to characterize a couple of actions and the quantity of causality for each of the ADRs pairs, correspondingly. $C < X, Y >$is evaluated by using the cue values represented in the table 1, where exact identification of cue values becomes the major important problem for each and every analyzes of pairs by corresponding matching of pairs with casual cue values

(e., very likely, probable, possible, and unlikely). In order to represent cue values we use a vector $V = (c_1, c_2, \ldots c_m)$values for each and every ADRs pairs for preprocessed data in the preprocessing stage. The experiences cue values results for each and every ADRs pairs is represented as $V = (c_1', c_2', \ldots c_m')$,then measure the similarity among those experiences and inexperienced cue values in the FRD model , it is denoted as $\text{fuzzy\_sim}(c_i, c_i')$,it values is represented inside the interval [0,1].

$$\text{fuzzy\_sim}(c_i, c_i') = \begin{cases} \text{poss}(c_i, c_i'), & \text{if poss}(\overline{c_i}, c_i') < 0.5 \\ \left(1.5 - \text{poss}(\overline{c_i}, c_i')\right) * \text{poss}(c_i, c_i'), & \text{otherwise} \end{cases}$$

Algorithm 1: Searching for drugs and support count for each drug

1. drughashtable = null
2. For each patient,$p_{it} \in$ DB do
3. Retrieve all the drugs D_{it} taken by the patient
4. For each drug $D_{it} \in$ DB do
5. if (drugHashTable.containsKey(d_{id} == false)do
6. $\sigma = 1$(a new adverse drug D_{txt} is found and set support count $\sigma = 1$)
7. else σ = drugHashTable.get Value (D_{k1}) + 1{update support count}
8. end if drugHashTable.put Value (D_{k1} , σ)
9. end for
10. end for
11. return drugHashTable

TABLE 1 : RELATING CUES TO CAUSALITY CATEGORY OF A SINGLE PAIR

Cues	Cue value set 1	Cue value set 2	Cue value set 3	Cue value set 4
Temporal association	Likely	Likely	Possible	Unlikely
Rechallenge	Likely	Possible	Unlikely	Unlikely
Dechallenge	Likely	Possible	Unlikely	Unlikely
Other explanations	No	No	No	Yes
Causality	Very likely	Probable	Possible	Unlikely

To optimize the parameters values in the FRD model in this paper presents an novel optimization algorithm to find best adverse drug reaction pairs in the system. In particle swarm optimization each and every particles is considered as fuzzy membership function to find best pairs .where each and every particles moves from one place to another place based on their velocity value and position are updated according to their membership function ,in general PSO algorithm finding the best pairs ,based on the cue values are updated according to (3-6),it optimizes the parameters values of the PSO algorithm by calculation of the weight values specifically in (5)(6) to improve the results of PSO algorithm so it is named as EPSO algorithm.

The general description of the EPSO for optimization of cue values in the membership function as follows

1. Initialize a population as number of cue values is considered as particles with random location and rapidity on D dimensions in the search space of best local cue values in the fuzzy membership function
2. loop
3. For each cue values from each pairs considered as particle, evaluate the desired optimization fitness function in D variables.
4. Compare particle's that is current pair cue values with its $pbestloc_i$ to find best cue values in the fuzzy membership function . If current value is better than $pbestloc_i$, then set $pbestloc_i$ equal to the current value of the best cue values for each pairs of the ADR
5. Initialize the particle's that is cue values with best known position to its initial position: $p_i \leftarrow x_i$
6. Identify the particles cue values in the neighborhood with the best ADR pair results and assign its index to the variable g.
   6.1. If ($f(p_i) < f(g)$) update the swarm's best known position: $g \leftarrow p_i$
7. Initialize the particle's velocity: $v_i \sim U(-|b_{upv}-b_{lov}|, |b_{upv}-b_{lov}|)$
8. Change the velocity and position of the particle cue values according to the following equation
9. Until a termination criterion is met ,repeat:
   9.1. For each particle (cue values) (i = 1, ..., S)
   9.2. do
   9.3. For each dimension d = 1, ..., n
   9.4. do
   9.5. Pick random numbers r_p, $r_g \sim U(0,1)$
   9.6. Update the particle's velocity $v_{i,d} \leftarrow \omega\ v_{i,d} + \varphi p_{rp}\ (p_{i,d}-x_{i,d}) + \varphi grg\ (g_d-x_{i,d})$

9.7. Calculate weight values using eq (5-6) is called as EPSO procedure

9.8. Update the particle's position with according to best cue values results for each pairs of the ADR pairs $x_i \leftarrow x_i + v_i$

9.9. If $(f(x_i) < f(p_i))$ do:

9.10. Update the particle's best cue values for each membership function position: $p_i \leftarrow x_i$

9.11. If $(f(p_i) < f(g))$ update the particles best features known position: $g \leftarrow p_i$

9.12. Now g holds the best found solution.

10. end loop.

Then now we concentrate to generate pairs with generated cue values before that the cue values in FRD model is optimized using EPSO algorithm. Based on the assignments of the velocity values only each and every particles move towards from best location to another best location. Once best optimized location are found then automatically the location of each and particles position are updated using position update formula, in this work we refer an EPSO. The each and every cue values in the FRD model is considered as particles ,it moves from one cue value to another cue values based on the velocity assignments for each particles ,and assessment of results to each use cue values in the FRD model [18-19]. The location of the cue values CFV the present location of the cue values particle , $v_i$the present velocity of the cue value by parameters v_{min} andv_{max} . The individual cue values of each and every pair is characterized by y_i. So the particle's location is familiar according to,

$$v_{cfv,k}(t+1) = wv_{cfv,k}(t) + c_1 r_{1,k}(t)\left(y_{cfv,k}(t) - x_{cfv,k}(t)\right) + c_1 r_{1,k}(t)\left(\hat{y}_k(t) - x_{cfv,k}(t)\right) \quad (1)$$

$$x_{cfv}(t+1) = x_{cfv}(t) + v_{cfv}(t+1) \quad (2)$$

Where w is the inertia weight whose series is [0-1], c_1 & c_2 are the randomly assignment best of values for each and every local and global cue values in the same way, cognitive factor and social factor, $r_{1,cfv}(t), r_{2,cfv}(t) \sim U(0,1)$ and $k = 1, \ldots, N_d$.

$$y_{cfv}(t+1) = \begin{cases} y_{cfv}(t) \text{ if } f(x_{cfv}(t+1)) \geq f(y_{cfv}(t)) \\ x_{cfv}(t+1) \text{ if } f(x_{cfv}(t+1)) < f(y_{cfv}(t)) \end{cases} \quad (3)$$

$$c_2 r_{2,k}(t)\left(\hat{y}_{cfv,k}(t) - x_{cfv,k}(t)\right) \quad (4)$$

Where $\hat{y}_{cfv,k}(t)$ are the most excellent cue values results for ADR pairs in the particle. The PSO is employ through recurrent compliance of the equation (1), (2) for all and every cue values of the each pairs in anticipation of a summing-up quantity of steps has been go beyond and equivalent to the rapidity of the all and every cue values in the FRPD that are most closer to maximum number of iterations. Once a novel g_{best} is recognized for every one of fuzzy cue values in the ADRs pairs in the health care records. In this improved cue values based optimization algorithm the fuzzy values are correctly optimized by consideration of different inertia values in the following manner ,

$$w = (w - 0.4)\left(\frac{MAXITER - ITERTATION}{MAXITER}\right) + 0.4 \quad (5)$$

Where MAXITER is considered as maximum number of iteration to complete best cue values identification in the ADR pairs represented in FRD models result and ITERATION characterizes the number of iterations to completed the cue values optimization process ,in order to speedup the process additionally add exponentially weight assignment of the above step and it is represented as Equ .(6)

$$w = (w - 0.4)e^{\left(\frac{MAXITER - ITERTATION}{MAXITER}\right)^{-1}} + 0.4 \quad (6)$$

The fitness function of this cue fuzzy membership fucntion is represented as follows :

$$fitness(P_{cfv}^d) = \sum_{i=1}^{swarm\_size} \frac{fuzzy\_sim(c_i, c_i'))}{swarm_{size}} \quad (7)$$

Algorithm 2 :Optimization of fuzzy membership function

1. Initialize every particle to contain cfvandomly assigned cue values
2. for t = 1 to t max do
3. for each particle cfvrandomly do

4. for each data vector cfv
5. The similarity of two position vectors in each and every cue value in the fuzzy membership function can be measured as $fuzzy\_sim(c_i, c_i')$,it values is represented inside the interval [0,1].
6. Assign fuzzy similarity to ADR pairs such that ,

$$fuzzy\_sim(c_i, c_i') = \begin{cases} poss(c_i, c_i'), & \text{if } poss(\overline{c_i}, c_i') < 0.5 \\ \left(1.5 - poss(\overline{c_i}, c_i')\right) * poss(c_i, c_i'), & \text{otherwise} \end{cases}$$

7. Calculate the fitness using equation (7)
8. Update the global best and local best positions
9. Update the cue value for FRD model with ADRs using (1) and (2)

From that cue values then mine the important cue values results for each and every ADRs pairs by measuring support and confidence values to each ADRs pairs in the health care records $(\text{support} > minsupp)$.To evaluate the strength of each and every ADRS pair for each patient records the simultaneously measure the results of both $< d_i, s_j >$ & $< s_i, d_i >$ usual causal leverage assessment, it is represented in the algorithm 3

Algorithm 3: Pair (Candidate Rule) Generation and evaluation

1. for each drug $d_i \in D$ do
2. for each symptom $s_j \in S$ do
3. retrievePID_s that support pair $< d_i, s_j >$ from database
4. if$(\text{count}(PID_s) \geq \text{mincount})$ then
5. value1 = casual-leverage (d_i, s_j, PID_s)
6. value2 = reverse casual-leverage (d_i, s_j) = casual-leverage (d_i, s_j, PID_s)
7. $\text{exclusive casual} - \text{leverage value} = \text{value}_1 = \text{value}_2$
8. Output pair $< d_i, s_j >$ and its exclusive casual-leverage value
9. end if
10. end for
11. end for

Similarly analysis the causal –leverage values for each ADR pairs in the dataset with exactly optimized cue values results in the EPSO algorithm, then these values are updated in the algorithm 4 ,it is greater than zero then is accepted as best mining results or else is considered as worst ADRs pairs in the FRD the causal leverage assessment of the couple is calculate and come back. In conclusion, we grade every one the pairs in a diminishing regulate corresponding to their restricted causal leverage ideals following each and every one these values are calculate.

Algorithm 4 Procedure casual –leverage(X,Y,PID_s)

1. search drug/symptom hash table to get support count for Y- σ_Y
2. for each PID that support the pair do
3. v=cue-abstraction(PID)
4. $SV = \{sv | sv = S_G(V, V') \cap V' \in EKB\}$(calculate the similarity values)
5. $SV' = normalization(SV)$
6. $C_{<X,Y>} = \text{weightedSum}(SV', W)$
7. if $C_{<X,Y>} > 0)$ then
8. accumulatedVotes+= $C_{<X,Y>}$
9. contributionCases + + {number of cases whose votes are greater than 0}
10. End if
11. End for
12. Supp$(X \xrightarrow{c} Y)$ =accumulatedVotes /N
13. Supp $(X \xrightarrow{c})$ =contributionCases /N
14. Supp$(\rightarrow Y) = \frac{\sigma_Y}{N}$
15. Return (Supp $(X \xrightarrow{c} Y)$-Supp $(X \xrightarrow{c})$ * Supp$(\rightarrow Y)$

4. EXPERIMENTATION RESULTS

In order to measure the performance of the proposed hybrid fuzzy recognition-primed decision (HRPD) model we collect dataset from Veterans Affairs Medical Center in Detroit, Michigan. Where the dataset are collected from knowledge database and stored in database that are related to electronic data that consist of all patients records with drug and symptoms were retrieved. For instance, the information for administration of drug comprise the surname of the drug, capacity of the drug distribute, amount of the medicine, medicine initiate time, drug list, and the amount of replenish.

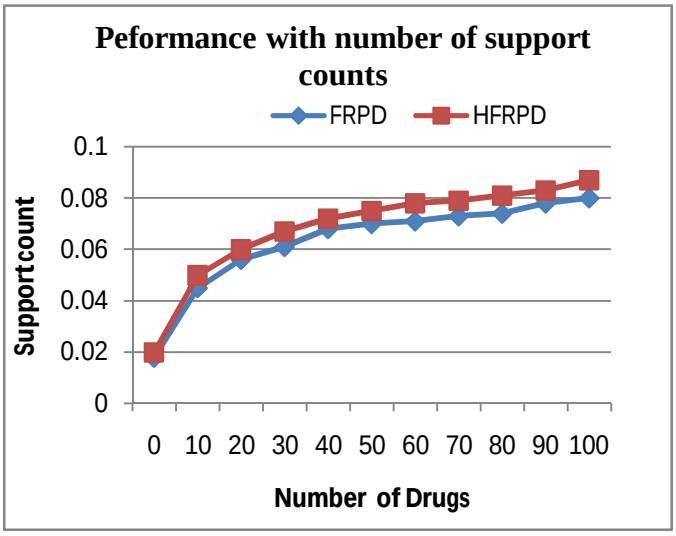

Figure 1: Performance comparison of fuzzy and HFRPD

In this graph we compare the performance of FRPD and proposed HFRPD system based on the Number of drugs the support count value changes at y axis .If the number of drugs is less and support count value also less, accordingly number of drug pairs increases as well as support value also high .Support count value of HFRPD is high than the FRPD system, is represented in Table 2.

Table 2: Performance comparison of FRPD and proposed HFRPD

Number of drugs	FRPD	HFRPD
	Support count	
0	0.018	0.02
10	0.045	0.05
20	0.056	0.06
30	0.061	0.067
40	0.068	0.072
50	0.07	0.075
60	0.071	0.078
70	0.073	0.079
80	0.074	0.081
90	0.078	0.083
100	0.08	0.087
110	0.081	0.088
120	0.084	0.091

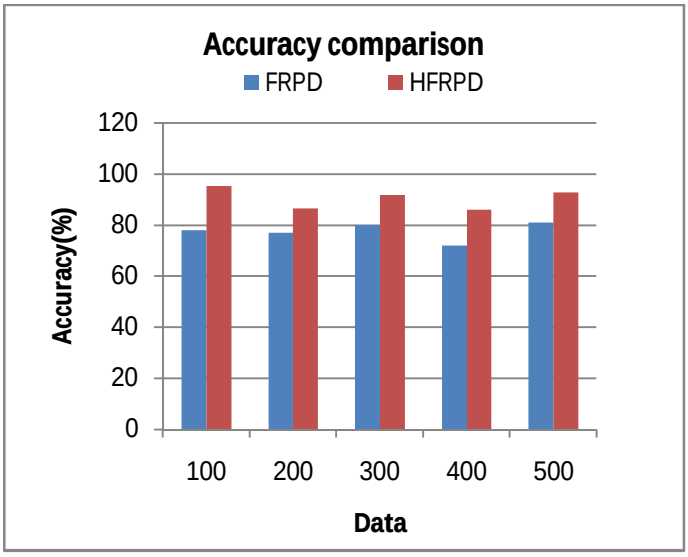

Figure 2: Accuracy comparison

In this graph we compare the accuracy of existing FRPD and proposed HFRPD system. Number of methods defines at X axis and percentage of accuracy at Y axis. Proposed HFRPD system achieves higher accuracy for finding the adverse drug pairs and optimized result than the Existing FRPD system, , is represented in Table 3.

Table 3: Accuracy comparison of FRPD and proposed HFRPD

Datasize(Kb)	FRPD	HFRPD
	Accuracy(%)	
100	78	95.5
200	77	86.5
300	80	92
400	72	86
500	81	93

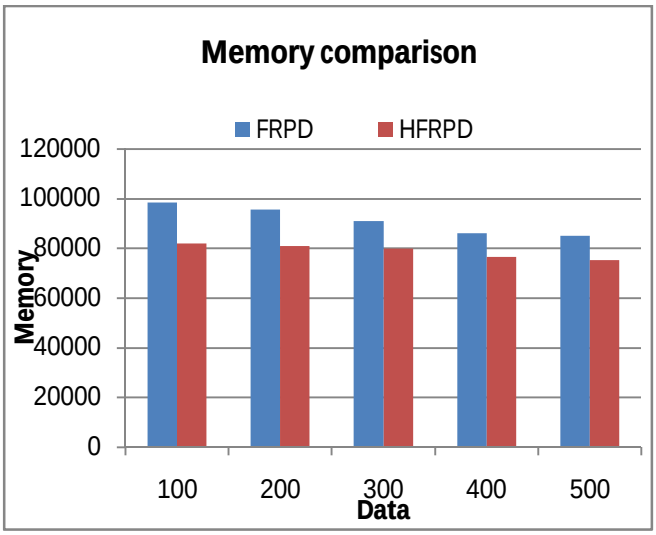

Figure 3: Memory comparison

In this graph we compare the memory taken to complete the existing FRPD and proposed HFRPD system. Number of methods defines at X axis and memory size to find drug pairs at a Y-axis is considered at bytes .Based on the system capacity proposed HFRPD system takes less memory capacity to discover the causal relationships can help us prevent or correct negative outcomes caused by its antecedents than the Existing FRPD based system, , is represented in Table 4.

Table 4: Memory comparison of FRPD and Proposed HFRPD

Datasize(Kb)	FRPD	HFRPD
	Memory	
100	98500	82000
200	95500	81000
300	91000	80000
400	86000	76500
500	85000	75250

5. CONCLUSION AND FUTURE WORK

In this paper proposed a hybrid FRPD method to mine exact casual relationship among drug and symptoms pairs among different patient in the preprocessed data from health care records. It can help to mine both causal and exclusive casual leverage analysis results using HFRD model ,it performs better than FRPD model ,since the parameters or cue values of the membership function is optimized correctly using EPSO algorithm ,it says that proposed system mine more ADRs pairs than fuzzy RPD model. Experimental results showed that our algorithm might successfully formulate identified ADRs rank elevated amongst each and every one the symptom in health care records .One of the important issues in the HFRPD model is that, it doesn't support for weighted association rule mining, to perform this initially create Lucene operator to support weighted association and mine weighted frequent item sets are introduced, finally perform same work in this paper.

REFERENCES

1. H.J. Murff, V.L. Patel, G. Hripcsak, and D.W. Bates, **"Detecting Adverse Events for Patient Safety Research: A Review of Current Methodologies,"** *J. Biomedical Informatics*, vol. 36, no. 1/2, pp. 131- 143, 2003.
2. P.E. Langton, G.J. Hankey, and J.W. Eikelboom, **"Cardiovascular Safety of Rofecoxib (Vioxx): Lessons Learned and Unanswered Questions,"** *The Medical J. Australia*, vol. 181, no. 10, pp. 524-525, 2004.
3. C. Kelman, S. Perason, R. Day, C. Holman, E. Kliewer, and D. Henry, **"Evaluating Medicines: Let's Use All the Evidence,"** *Medical J. Australia*, vol. 186, no. 5, pp. 249-252, Mar. 2007.
4. J. Chen, H. He, G. Williams, and H. Jin, **"Temporal Sequence Associations for Rare Events,"** *Proc. Pacific-Asia Conf. Knowledge Discovery and Data Mining (PAKDD '04)*, pp. 235-239, May 2004.
5. C.L. Burgess, C.D. Holman, and A.G. Satti, **"Adverse Drug Reactions in Older Australians, 1981-2002,"** *The Medical J. Australia*, vol. 182, no. 6, pp. 267-270, Mar. 2005.
6. M. Hauben and X. Zhou, **"Quantitative Methods in Pharmacovigilance: Focus on Signal Detection,"** *Drug Safety*, vol. 26, no. 3, pp. 159-186, 2003.
7. V. Curcin, M. Ghanem, M. Molokhia, Y. Guo, and J. Darlington, **"Mining Adverse Drug Reactions with E-science Workflows,"** *Proc. IEEE Cairo Int'l Biomedical Eng. Conf. (CIBEC '08)*, pp. 1-5, Dec. 2008.
8. D.W. Bates, R.S. Evans, H. Murff, P.D. Stetson, L. Pizziferri, and G. Hripcsak, **"Detecting Adverse Events Using Information Technology,"** *J. Am. Medical Informatics Assoc.*, vol. 10, no. 2, pp. 115-128, 2003.
9. J. Li, A.W.-C. Fu, H. He, J. Chen, H. Jin, D. McAullay, G. Williams, R. Sparks, and C. Kelman, **"Mining Risk Patterns in Medical Data,"** *Proc. ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD '05)*, pp. 770-775, 2005.
10. M. Stephens, J. Talbot, and P. Routledge**," Detection of New Adverse Drug Reactions",** *Macmillan Reference Ltd*, 1998.
11. Kum, H.-C., Duncan, D., Flair, K. & Wang, W. (2003**)," Social welfare program administration and evaluation and policy analysis using knowledge discovery and data mining (kdd) on administrative data, in",** '*Proceedings of the NSF National Conference on Digital Government Research (DG.O)*', pp. 39–44.
12. Murff, H. J., Patel, V. L., Hripcsak, G. & Bates, D. W. (2003), "**Detecting adverse events for patient safety research: a review of current methodologies"**, *Journal of Biomedical Informatics* 36(1/2), 131–143.
13. Harvey, J., Turville, C. & Barty, S. (2004), "**Data mining of the australian adverse drug reactions database: a comparison of bayesian and other statistical indicators"**, *International Transactions in Operational Research* 11(4), 419–433.
14. Chen, J., He, H., Williams, G. & Jin, H. (2004), **"Temporal sequence associations for rare events",** in '*Proceedings of 8th Pacific-Asia International Conference on Knowledge Discovery and Data Mining (PAKDD)*, Lecture Notes in Computer Science (LNAI 3056)', Sydney, Australia, pp. 235–239.

15. Almenoff, J. S., DuMouchel, W., Kindman, L. A., Yang, X. & Fram, D. (2003), **"Disproportionality analysis using empirical bayes data mining: a tool for the evaluation of drug interactions in the post-marketing setting",** *Pharmacoepidemiology and Drug Safety* 12(6), 517–521.

16. Parsopoulos, K. E. & Vrahatis, M. N. (2001). **"Particle swarm optimizer in noisy and continuously changing environments".** *InM. H. Hamza (Ed.), Artificial intelligence and soft computing. Anaheim: IASTED/ACTA.* pp. 289–294

17. Pugh, J., Martinoli, A. & Zhang, Y. (2005**)." Particle swarm optimization for unsupervised robotic learning."** *In Proceedings of IEEE Swarm Intelligence Symposium (SIS),* pp. 92–99), Piscataway: IEEE.

18. Cui, X., Potok, T., Palathingal, P., **"Document Clustering using Particle Swarm Optimization",** *Swarm Intelligence Symposium, 2005. Proceeding*s 2005 IEEE, pp. 185- 191.

19. Li-ping, Z., Huan-jun, Y., Shang-xu, H., **"Optimal Choice of Parameters for Particle Swarm Optimization**", *Journal of Zhejiang University Science*, Vol. 6(A)6, pp.528-534, 2004.