transaction for the same user based on their balance we can easily find out reconcile statement among that id.

a) Greedy Hierarchical Item Set-Based Clustering (GHIC) Algorithm

The best clustering algorithm is used for the process of grouping the transactions which will be easy for the process of reconciling the required data [4]. After obtaining the FIS, set of frequent access item sets we are creating a variable C0 which contain all the transactions. These two variables C0 and FIS are obtained as inputs and finally the output is obtained on the variable C which symbolically represents the cluster [5], [6], and [7].

Algorithm:

```
Inputs: $C_0=\{T_1,T_2,\ldots\ldots T_n\}$
        $FIS=\{is_1,is_2,\ldots\ldots is_m\}$
Output: Cluster $C=\{C_1,C_2,\ldots\ldots C_j\}$
            $C=\{C_0\}$
        Repeat{
        Choose any $X=\{T_1,T_2,\ldots\ldots T_n\}$
        For each isr in FIS{
            Let $C_a=\{T_k|T_k$ € X and $is_r \subseteq T_k\}$;
            Let $C_b=X-C_a$
            Calculate Mr=M(ca,cb)
            }
            C=(C-X)U {Ca}U{Cb} ;
            }
```

Then the required specific transaction is stored on the variable X then the specified function is carried on and as the final output the clustered data are obtained. Which is used for the further process? Figure 1 explains the algorithm that is carried on GHIC.

b) Elucidation of System Architecture

As the critical reconciliation process between the two organizational accounts are explained clearly with the help of the Figure 2. In this diagram the input from the organization is obtained and then then are taken for the reconciliation process and here the data are extracted using the GHIC algorithm and they are separated based on their transaction system such as NEFT or RTGS. After considering the required transaction scheme the pattern visualization is carried on. In the pattern visualization process they will display the kind of transaction which is considered for the process.

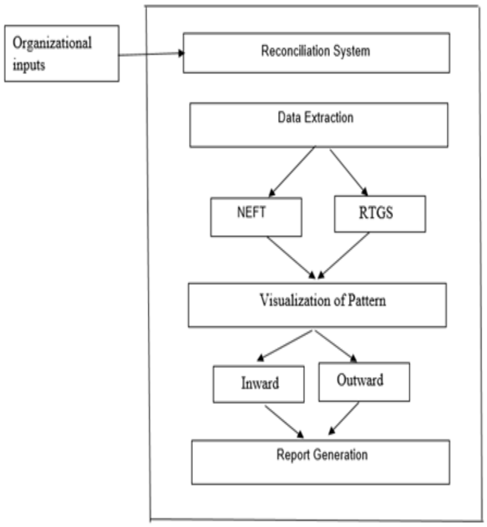

Figure 2: System Architecture

In further process they will check the type on transaction such as Inward (amount obtained from other organization) or Outward (amount withdrawn by other organization) which are carried out and also they will compare the status of the two files using the primary key [8],[9]. If the amount present in the two organizational accounts are same then they will not display the transaction id else if there is a difference between the amount deposited and the amount withdrawn then they will highlight the transaction id indicating that the transaction have not completed properly[10]. Thus the status of the transaction is highlighted as the final result.

III. RESULTS AND DISCUSSION

In the final display screen the result will be displayed. In the result the transaction which doesn't have the proper reconciliation will be highlighted indicating that there is an occurrence of improper transaction [11]. From this we can easy find the odd transactions from the huge number of available transactions of the particular organization. The major advantage is that they can easily collect the transactions information by the clustering order and thus they can be grouped easily under the required criteria. As the future enhancement of the subject we can have a reconciliation checking for time to time and thus we can reduce the waiting time of the result and increase the accuracy of the project.

IV. CONCLUSION

As discussed in this paper, the existence of the multiple transactions reconciliation result can be obtained quickly and accurately. We suggest that the Hierarchical Clustering technique will be suitable for the process of grouping the transaction according to the requirements. In this paper, we presented a new approach, GHIC (Greedy Hierarchical Item Set-Based Clustering) for pattern-based clustering of organizational transactions and demonstrated that the technique performs effectively, compared to some traditional techniques. Further enhancement can also create great advantages to this current process.

V. REFERENCES

[1] Hofner. P, “Algebraic View Reconciliation”, Published in: Software Engineering and Formal Methods, 2008. SEFM '08. Sixth IEEE International Conference on 10-14 Nov. 2008.

[2] B. E. Jacobs. Applied Database Logic: Fundamental Issues, volume I. Prentice-Hall, Inc., 1985.

[3] R. Khedri. Formal model driven approach to deal with requirements volatility. Computing and Software Technical Reports CAS-08-03-RK, Department of Computing and Software, McMaster University, 2008.

[4] Yinghui Yang and Balaji Padmanabhan, GHIC: A Hierarchical Pattern-Based Clustering Algorithm for Grouping Web Transactions, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, SEPTEMBER 2005.

[5] M. Steinbach, G. Karypis, and V. Kumar, “A Comparison of Document Clustering Techniques,” Proc. Int’l Conf. Knowledge Discovery and Data Mining Workshop Text Mining, 2000.

[6] K. Wang, C. Xu, and B. Liu, “Clustering Transactions Using Large Items,” Proc. ACM Int’l Conf. Information and Knowledge Management (CIKM ’99), pp. 483-490, 1999.

[7] S. Kimbrough, B. Padmanabhan, and Z. Zheng, “On Usage Metric for Determining Authoritative Sites,” Proc. Workshop Information Technology & Systems (WITS 2000), pp. 43-48, 2000.

[8] E. Han, G. Karypis, V. Kumar, and B. Mobasher, “Clustering Based on Association Rule Hypergraphs,” Technical Report TR-97-019, Dept. of Computer Science, Univ. of Minnesota, Minneapolis, 1997.

[9] G. Karypis, R. Aggarwal, V. Kumar, and S. Shekhar, “Multilevel Hypergraph Partitioning: Application in VLSI Domain,” Proc. 34th Ann. ACM/IEEE Design Automation Conf., pp. 526-529, 1997.

[10] H. Wang, J. Yang, W. Wang, and P.S. Yu, “Clustering by Pattern Similarity in Large Data Sets,” Proc. ACM SIGMOD 2002 Conf., pp. 394-405, 2002.

[11] Y. Yang, X. Guan, and J. You, “CLOPE: A Fast and Effective Clustering Algorithm for Transactional Data,” Proc. ACM Int’l Conf. Knowledge Discovery and Data Mining (KDD ’02), pp. 682-687, 2002.