# Trust network based prediction for user ratings in trip advisor

Anushka Harshad Patil
*Computer and Information Sciences*
Indiana University Purdue University Indianapolis
Indianapolis, Indiana
ahpatil@iu.edu

Snehal Suresh Vyawahare
*Computer and Information Sciences*
Indiana University Purdue University Indianapolis
Indianapolis, Indiana
svyawaha@iu.edu

*Abstract*

**Now-a-days, recommender systems are used by a wide variety of users. It aims to provide a trusted review about different businesses. In this project we have used a trust system that evaluates the trust based on the author's interaction and relationship with other authors. This system collects and maintains data about different authors, their reviews and hotel details and estimates a relation between them. It determines trust based on the ratings given of users given to each other. This relationship is given in terms of reviews and ratings on their reviews. The author's reputation (trust) is computed based on its opinion, rating scores and review influence.**

*Keywords— E-Commerce; Trust Management; Ratings; Trip Advisor*

## I. INTRODUCTION

Precise trust evaluation is vital for the success of systems that work on the user reviews. User reviews are recognized as the sales drivers and it is noticed that most customers will want to analyze the reviews before deciding to make purchase or to selecting any online service. It is seen that the number of online users who read and trust online reviews is increasing and thus, it is essential to provide mechanism that will check the quality of the review in terms of how trustworthy the review is. The customer reviews on any site plays an important in establishing trust with the new customers and converting that to an increased sale.

In this project, we have analyzed the Trip Advisor site. TripAdvisor, Inc. is an American travel and restaurant website company providing hotel and restaurant reviews, accommodation bookings and other travel-related content. Trip Advisor is based on the user generated content and provides free services to users. The website is supported by a hotel booking facility and an advertising business model [6]. This project we derive trust information from other user ratings given to the users for hotel bookings. We have used trust management framework and adapt

it to Trip advisor dataset. Also, we use simple methods to calculate the confidence between the users to determine the trust and errors which may be present. By applying the formula to predict the ratings which users have given to each other.

## II. LITERATURE SURVEY

### A. A General Trust Management Framework for Internet of Things

Managing trust relationship amongst different agents and their partners is a challenging task and a very important topic for Internet of Things. The paper "*A General Trust Management Framework for Internet of Things*" by *Yefeng Ruan, Arjan Durresi, Lina Alfantoukh*. [4] proposed a trust management framework that will help agents to evaluate their partners' trustworthiness. The focus of this paper is to provide a general framework irrespective of determining the environments or factors for trust in a specific application. In this framework they have represented trust using Trustworthiness and confidence. Trustworthiness is the comprehensive summary of the subject's multiple measurements or evaluations towards the object, while confidence measures to what extent the subject is certain about the trustworthiness evaluation. Following are the Trust modeling schemas from the paper
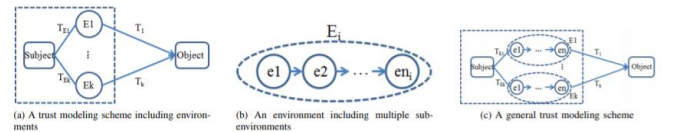


Fig. 1. *Trust modeling schemas*

To illustrate the usage of framework they have performed an experimental simulation for a food nutrition analysis in diabetes treatment by assuming that there are 100 patients who are looking for services. Here, these patients themselves are acting as tagers, identifiers and measures. They have assumed patients are from two groups based on their food types. The results showed that by using trust it can help patients to filter out inaccurate information or mitigate its effect. Additionally, they have illustrated the attacks stimulation to demonstrate the importance of selecting and weighting environments in IoT systems.

### B. Trust Assessment for Internet of Things in Multi-access Edge Computing.

[1] Multi- access edge computing aims to improve the delivery of content and applications to users. Its main aim is to provide end users with a network which is congestion free and improve application performance. This, paper describes about use of micro clouds. It is due its good performance, multiple-access edge computing is considered as a platform for various applications such as augmented reality, e-governance, e-health, gaming, IOT. MEC contains various micro clouds like public, private and hybrid. This paper proposes to design a strong security architecture for multiple-access edge cloud computing - IoT systems. This paper has adapted and applied the measurement theory based on trust management framework for large infrastructures. This paper makes use of trustworthiness along with confidence. It states the trust relation and methods of trust evaluation in MES.

Trust management framework-

It has two metrics- trustworthiness and confidence. Trustworthiness (m) is represented as a comprehensive summary of multiple "measurements". It measures to what extent the user trusts the trustee. It shows to what extent we are confident about the trustworthiness evaluation. It represents the variance of the actual value from the summarized value. Therefore, second parameter is taken into consideration i.e. confidence(c).

$$m = \frac{\sum_{i=1}^{i=k} m_i}{k} \quad \ldots(a)$$

$$C = \begin{cases} 1 - 2 * r, & if\ r \leq 0.5 \\ 0, & otherwise \end{cases} \quad \ldots(b)$$

### III.  METHODOLOGY

In ecommerce sites, user reviews are recognized as the sales drivers and it is noticed that most customers will want to analyze the reviews before deciding to make purchase or to select any online service.
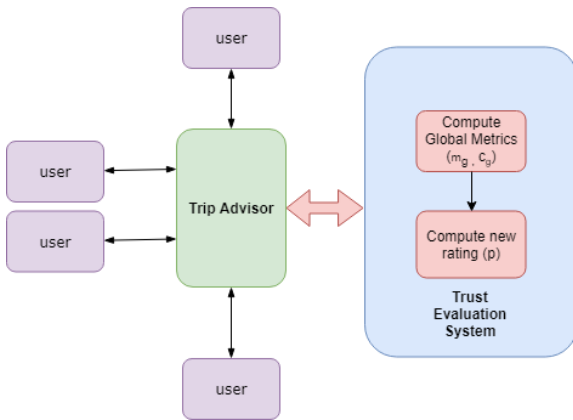


Fig. 2.  *Trust-based Trip Advisor System*

It is seen that the number of online users who read and trust online reviews is increasing and thus, it is essential to provide mechanism that will check the quality of the review in terms of how trustworthy the review is. The customer reviews on any site plays an important in establishing trust with the new customers and converting that to an increased sale.

### A.  Data

For this project, we have proposed a methodology that will use the trust between different authors to compute the trust-based rating for hotels. We have created some dummy datasets to illustrate the working of our system.

### B.  Global metrics

Our trust-based Trip Advisor system, uses these dummy datasets to compute the global metrics: Trustworthiness and confidence. To compute these global metrics, we have aggregated the author opinions, rating scores and review influence on other authors. The authors will rate the hotel reviews by other authors and this relationship in terms of reviews and ratings on their reviews will be used to evaluate the global trustworthiness 'mg' and global confidence 'cg'

$$m_g = \frac{\sum i \in N \frac{1}{\sigma_i^2} * m_i}{\sum i \in N \frac{1}{\sigma_i^2}} \quad \ldots(c)$$

$$c_g = 1 - 2 * \sqrt{\frac{1}{\sum i \in N\ \sigma_i^2}} \quad \ldots(d)$$

On calculating these global metrics, we apply them on our trust-based system to compute the new ratings based on the interactions, confidence and variance of an author and evaluating it with the author opinions, rating scores and review influence about the hotel.

### C.  New Rating

To compute the new rating 'p' we will be using the following formula

$$p = \frac{\sum_1^N m_i * c_i * r_i}{\sum_1^N m_i * c_i} \quad \ldots(e)$$

The new rating is based on the aggregation of products of global metrics along with individual rating by each author from the author list for a hotel. Each author has his own global metrics based on his/her interactions with other authors and the review influences on the authors for any hotel.

### IV.  IMPLEMENTATION

For implementing the above methodology, we are using Java in IntelliJ IDE.

### A.  Dataset

We have created the dummy datasets in JSON format. Following are some of the snippets of the JSON file.

*1)  AuthorInfo.json:* This JSON file has information about the Author and has an Author Reviewer List which contains

the name of authors that have rated the author. This file also contains those ratings, interactions and variance for the author.

*2) Hotel_Info.json:* This JSON file has information about the Hotel and has an Author List which contains the name of authors that have rated the hotel. This file also contains the overall rating given by the above authors and other detailed information about the hotel.

*3) Reviews.json:* This JSON file has information about the reviews. It contains individual review by authors on a Hotel. It contains the detailed information about each review for a hotel.

### B. Code snippet

Following is code snippets for computing the new Rating based on trust model.

```java
private static void calculateNewRating() {
    JSONObject obj1 = (JSONObject) hotelsInfo;
    Double num = 0.0;
    Double deno = 0.0;
    JSONArray hArray = (JSONArray) obj1.get("Hotel");
    Double[] globalValue;
    System.out.println("HotelId    oldRating      newRating");
    for (Object h : hArray) {
        JSONObject hotel = (JSONObject) h;
        Double oldRating = (Double) hotel.get("Rating");
        String id = hotel.get("HotelID").toString();
        int n = Integer.parseInt(hotel.get("HotelRatingCount").toString());
        JSONArray r = (JSONArray) hotel.get("AuthorList");
        String[] AuthorList = new String[n];
        int j = 0;
        for (Object d : r) {
            AuthorList[j] = d.toString();
            j++;
        }
        JSONObject obj2 = (JSONObject) authorInfo;
        JSONArray aArray = (JSONArray) obj2.get("AuthorInfo");
        JSONObject obj3 = (JSONObject) reviews;
        JSONArray rArray = (JSONArray) obj3.get("Reviews");
        for(Object a : aArray){
            JSONObject author = (JSONObject) a;
            for(int i=0; i<AuthorList.length; i++) {
                if (author.get("Author").equals(AuthorList[i])) {
                    globalValue = AuthorGlobalValues.get(author.get("Author"));
                    for(Object re : rArray ){
                        JSONObject review = (JSONObject) re;
                        if(review.get("Author").toString().equalsIgnoreCase(AuthorList[i])){
                            num += globalValue[0]* globalValue[1]* (double) review.get("Ratings");
                            deno += globalValue[0]* globalValue[1];
                            break;
                        }
                    }
                }
            }
        }
    }
    Double newRating = num/deno;
    System.out.println(id+"     "+oldRating+"         "+df2.format(newRating));
}
```

Fig. 3. *Code snippet for computing the new rating*

Here, we have first parsed the JSON files into object and then used them to compute the global metrics: Trust and Confidence. Then we have created a function that computes the new rating based on the global metrics, author ratings and the overall rating of the hotel. The code snippet above shows this new rating function. Thus, using Java we were able to implement our methodology into practice.

## V. RESULTS AND DISCUSSION

To evaluate the relationship between the authors and to computer the global metrics, we created a dummy dataset which appears like the Trip Advisor dataset. To have sufficient data to estimate the results, we created datasets for 50 Authors, 50 Reviews and 10 hotels. On average each author rates at least 2-3 hotels and each author have an author reviewer list containing at least 3 authors. Our major focus

here was to compute the global metrics based on the interactions, author ratings and variance.

For experimenting, we created two dummy datasets.

*1) Random dataset:* These datasets have values put on random there is not modification done between genuine authors (good users) and fake authors (bad users)

*2) Classified dataset:* These datasets have values classified into genuine authors (good users) and fake authors (bad users).

On running these datasets over our trust-based trip advisor model, we were able to compute the new ratings. To illustrate this, we experimented with both the datasets and compared it with the old ratings and found that there were notable changes between the old ratings and new ratings. The old rating was based on the average of all rating irrespective of the trustworthiness of those ratings. The new ratings on the other hand were based on the trust i.e. the aggregate of the author opinions, rating scores and review influence on other authors.

On using *Random dataset,* we found that the old ratings and new ratings have a notable change and varies randomly. For certain hotel the ratings seemed to be increased while for some our trust model reduced the new rating. Fig .4. shows the detailed comparison between the old ratings and the new ratings. We have plotted the graph for the random datasets for better analysis Fig .5.



Fig. 4. *Comparision between the old ratings and new ratings for each hotel*
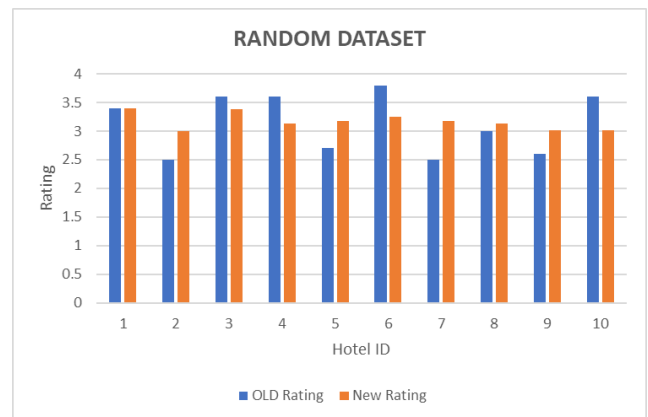


Fig. 5. *Graph for analysing the change in rating on random dataset*

On using *Classified dataset,* we classified the authors as genuine and fake such that author id from 1 to 50 are in order of genuine to fake. The global metrics will give a detailed picture of the exact classification. Fig.6.

We have considered the scenario wherein, genuine users rate hotel_01 with genuine reviews and the new ratings is computed based on it and it shows how precisely the model works. Second scenario, fake users have rated Hotel_10 with excellent rating wherein Hotel_10 doesn't deserver that many ratings. The other authors have thus rated these fake users with lower rating for their fake reviews. Our trust model, examines these factors and reduces the overall rating of the hotel for reliable and precise rating.

| AuthorId | Global m | Global c | | AuthorId | | |
|----------|----------|----------|---|----------|-----|------|
| Author01 | 5.0 | .31 | | Author25 | 3.0 | .21 |
| Author02 | 5.0 | .33 | | Author26 | 3.0 | .2 |
| Author03 | 5.0 | .24 | | Author27 | 3.0 | .24 |
| Author04 | 5.0 | .34 | | Author28 | 3.0 | .22 |
| Author05 | 5.0 | .25 | | Author29 | 3.0 | .2 |
| Author06 | 5.0 | .33 | | Author30 | 3.0 | .22 |
| Author07 | 5.0 | .31 | | Author31 | 2.0 | .2 |
| Author08 | 5.0 | .21 | | Author32 | 2.0 | .2 |
| Author09 | 5.0 | .32 | | Author33 | 2.0 | .22 |
| Author10 | 5.0 | .24 | | Author34 | 2.0 | .21 |
| Author11 | 4.0 | .23 | | Author35 | 2.0 | .18 |
| Author12 | 4.0 | .26 | | Author36 | 2.0 | .21 |
| Author13 | 4.0 | .19 | | Author37 | 2.0 | .25 |
| Author14 | 4.0 | .23 | | Author38 | 2.0 | .27 |
| Author15 | 4.0 | .21 | | Author39 | 2.0 | .27 |
| Author16 | 4.0 | .18 | | Author40 | 2.0 | .19 |
| Author17 | 4.0 | .27 | | Author41 | 1.0 | .24 |
| Author18 | 4.0 | .21 | | Author42 | 1.0 | .24 |
| Author19 | 4.0 | .23 | | Author43 | 1.0 | .21 |
| Author20 | 4.0 | .19 | | Author44 | 1.0 | .31 |
| Author21 | 3.0 | .21 | | Author45 | 1.0 | .35 |
| Author22 | 3.0 | .21 | | Author46 | 1.0 | .22 |
| Author23 | 3.0 | .2 | | Author47 | 1.0 | .27 |
| Author24 | 3.0 | .21 | | Author48 | 1.0 | .25 |
| Author25 | 3.0 | .21 | | Author49 | 1.0 | .18 |
|  |  |  | | Author50 | 1.0 | .21 |

Fig. 6. *The global trust metrics:global m- trustworthiness and global c-confidence*

Fig .7. shows the detailed comparison between the old ratings and the new ratings. We have plotted the graph for the classified datasets for better analysis Fig .8

| HotelId | oldRating | newRating |
|---------|-----------|-----------|
| Hotel_01 | 4.9 | 4.9 |
| Hotel_02 | 1.2 | 3.5 |
| Hotel_03 | 3.0 | 3.27 |
| Hotel_04 | 1.0 | 2.7 |
| Hotel_05 | 1.83 | 2.41 |
| Hotel_06 | 4.7 | 2.71 |
| Hotel_07 | 4.8 | 2.95 |
| Hotel_08 | 2.0 | 2.87 |
| Hotel_09 | 2.6 | 2.71 |
| Hotel_10 | 4.6 | 2.91 |

Process finished with exit code 0

Fig. 7. *Comparision between the old ratings and new ratings for each hotel*



Fig. 8. *Graph for analysing the change in rating on classified dataset*

For classified dataset, we found that the old ratings and new ratings have a notable change and varies accurately based on the global trust metrics. The new ratings seemed more reliable, précised and trustworthy.

This analysis indicates that our methodology works well for computing the hotel ratings based on the author opinions, rating scores and review influence on other authors rather than just calculating mean for all the ratings.

## VI. Conclusion

In this project, we used the dummy datasets and a trust management framework for our trust-based trip advisor system. Based on the author-author interaction and confidence we were able to compute the new ratings using our trust-based system. To see whether this system makes any notable change in the original trip advisor system, we compared the new ratings with the old ones. The new ratings seemed to be more prominent and trustworthy. Thus, the result analysis above clearly shows that by using our trust-based system to compute the ratings we can have an improved, précised and reliable system.

## VII. Future work

Our work has some limitations. Firstly, we performed experimental analysis over small amount of dataset i.e. for relatively less number of authors, hotels and reviews. It is possible that on increasing the number of datasets or varying data might result into something new. Therefore, our future work focuses on testing this model over different datasets. Additionally, our model as of now focuses on the trip advisor system, we intend to integrate this methodology over other ecommerce sites which rely on the user opinions and reviews. Thus, our future work also focuses on integrating our model on different ecommerce sites.

## REFERENCES

[1] Yefeng Ruan, Arjan Durresi, Suleyman Uslu, "Trust Assessment for Internet of Things in Multi-access Edge Computing"

[2] Yefeng Ruan, Ping Zhang, Lina Alfantoukh, and Arjan Durresi, "Measurement Theory-Based Trust Management Framework for Online Social Communities"

[3] Yefeng Ruan, Arjan Durresi, Lina Alfantoukh, "Using Twitter trust network for stock market analysis"

[4] Yefeng Ruan, Arjan Durresi, Lina Alfantoukh "A General Trust Management Framework for Internet of Things"

[5] Kushagra Bhargava, Tanvi Gujral, Mihir Chawla , Tanya Gujral, "Comment Based Seller Trust Model for E-Commerce"

[6] Wikipedia