User Trust and Malicious Voting

Western Michigan University

Jonah Kubath

**Crowd Sourced Data**

Most of today’s search engines rely on the use of off-loading work to those that use their product. Crowd sourcing data is when a collective group of people or users come together to combine ideas, opinions, or reviews to form a collective opinion on a “product”. The term “product” here can relate to businesses, items for sale, or honest and dishonest files in P2P file sharing. The users in the crowd can vote on the products. Their votes are then summed and published to the rest of the crowd so other users can make decisions based on other’s opinions. Crowd sourcing data has a few benefits. The first is that the search engine does not have to create new information on its own. With incentive for users, the users will add their own opinions on products to make more information available to others. The second is that users can be more confident in the opinion of a product when it is the aggregate of many users. Many users decrease the change of biased or even malicious data to be the group’s opinion on a product.

As hinted to in the second benefit of crowd sourced data, there is an opportunity for malicious users to change or impact group’s opinion on a product. One of the best solutions to prevent bad information from malicious users is to assign trust scores to all the users in the system. This will put more trust in users who have a record of voting honestly on other businesses and less trust in new users or users who previously voted dishonestly on other businesses. Voting honestly is determined when a user’s vote matches the aggregate of all the user’s votes on a specific product. As outlined in “Voting Systems with Trust Mechanisms in Cyberspace: Vulnerabilities and Defenses” by Feng, Sun, Liu, Yang, and Dai, user trust scores can be attached efficiently with their algorithm called Reputation Trap (RepTrap).

**Description of RepTrap**

The overall summary of RepTrap is a coordinated attack by the malicious users. The malicious users use the review information to determine a set of businesses that were also voted on by the users who voted on the target business called the correlated set. They will then collectively vote against the crowd on the correlated business set to reduce the system’s trust in other users while simultaneously increasing the system’s trust in the malicious users. This algorithm will reduce the total number of votes needed to attack the target business.

The first step in RepTrap is to calculate the correlated set and uncorrelated set of businesses. The correlated set is defined by any business “A” in which a user has voted on the target business and business “A”. The uncorrelated set is any business not in the correlated set.

The second step is to attempt to trap the target business. Trapping a business will include having all the malicious users vote on the target business. In application this can be changing a business’ review from good to bad, product stars from five to one, or an honest file to a dishonest file in P2P file sharing. If the malicious users do not have enough trust to trap the target business, RepTrap moves on to the next step.

The third step is to trap businesses in the correlated set. Changing the opinion on a correlated business will decrease the system’s trust in the honest users who voted on that business while increasing the system’s trust in the malicious users. The second step would then be attempted again after each business is trapped.

When all the businesses in the correlated set have been trapped, the malicious users will then attempt to trap uncorrelated businesses. While this will no longer decrease the trust of the users who voted on the target business, it will continue to increase the system’s trust in the malicious users. As before, the second step would be attempted after each business is trapped.

If the third and fourth steps fail to increase the malicious user’s trust enough to trap the target business, the malicious users will revert to voting honestly on businesses. They will start with businesses in the uncorrelated set as it will not increase trust in the users who voted on the target business and finish with the correlated set if needed. Once the process of voting honestly on a business is finished, trapping of the target business will be attempted in step two. If this fails, the number of malicious users was not great enough to attack the target business in the system.

**Application of user trust and malicious voting**

The data used in this project was publicly available data from yelp.com. The dataset includes business information and reviews from users. The yelp dataset was used as a real-world implementation of the RepTrap attack method. Users are assigned trust scores based on the number of aggregate crowd votes that match the vote given by the user. The number of matching votes will be called “good” and the number of votes against the crowd will be “bad”. The total user trust is: (good + 1) / (good + bad + 2). This defaults all new users to a trust score of 0.50 and an upper limit of less than 1.0. In the yelp dataset, the users can give star ratings from one to five. To simplify computation, these values are converted to binary as anything below three stars is a bad review and three and above are good reviews. The final business stars are calculated by dividing the summation of user’s trust who voted a business good divided by the summation of all user’s trust who voted on the given business.

**Optimizations to RepTrap Algorithm**

During the study and implementation of the RepTrap algorithm, an optimization was found to decrease the total number of votes required by the malicious users. The general RepTrap algorithm is to trap all the correlated businesses, then trap uncorrelated, and finally vote honestly on the uncorrelated and correlated sets. The optimization that was implemented was to calculate the theoretical number of honest votes required to raise the total malicious trust score to a high enough value that allows the target business to be trapped. This allows the RepTrap algorithm to break from trapping businesses in the correlated or uncorrelated sets early, when it would be beneficial (require fewer malicious votes) to simply vote honestly on a few uncorrelated businesses. The results from this optimization will be discussed in the next section.

**Figure 1** is the total number of votes required to trap a target in the top 10 percent of the businesses with respect to positive reviews.

**Figure 2** is the number of votes required by each attack strategy.

**Figure 3** is the total number of votes required to trap a target in the top 1.6 percent of the businesses with respect to positive reviews.

**Figure 4** is the number of votes required by each attack strategy.

**Results from RepTrap**

The RepTrap algorithm was implemented on the public yelp dataset. The dataset contained 6,185,900 individual reviews for 192,609 different businesses. The review data is read in and the simulation data is updated. The simulation data contains business stars and the trust of each user. The business stars are calculated by the ratio: summation of user’s trust who voted positively on the business divided by the total trust of all users who voted on the business. This ratio is then multiplied by a max value of five stars. The trust of each user is calculated by: (number of the user’s votes that match the crowd + 1) / (total user’s votes + 2). Once this information is updated, the number of malicious users indicated by the user are created and the RepTrap is started. Various business ids were selected as the target business. As would be expected, as the number of malicious user ids increases, in Figure 1 and 3, the total number of votes required to trap the target decreases. The businesses were sorted by the total summation of user’s trust scores that voted three or more stars on the business. The total trust needed to trap the top %10 business was 40.49 and to trap the top %1.6 business was 183.95. When malicious users only vote honestly to increase their trust score, a trust score of approximately 1.0 is the highest possible score. This means that whenever the number of malicious users is less than the ceiling of the total business trust, the trap will fail when only honest votes are used. While testing the simulation, RepTrap and the optimized version of RepTrap were able to continue trapping the target business with %25 fewer resources than required by only voting honestly. In Figure 1, voting honestly needed 41 malicious users to trap the target, but RepTrap was able to trap the target business with 34 malicious users. In Figure 3, voting honestly needed 190 malicious users to trap the target, but RepTrap was able to trap the target business with 180 malicious users. The 180 malicious users could have been reduced, but the total number of malicious votes would drastically increase.

**Results from Optimization**

The optimization that was added to the RepTrap algorithm was the ability to bypass the trapping of a correlated or uncorrelated business and go directly to voting honestly. When this is done in the final stages of the RepTrap algorithm, the total amount of trust needed to be gained by the malicious users is very small. As an example, the next best business to trap may require 10 malicious votes to trap, but only 4 honest votes are required to gain enough malicious trust to trap the target business. This difference (10 votes – 4 votes = 6 votes saved) is the optimal voting strategy. On paper, the difference in voting strategies is only marginally different. In practice, the more votes used by the malicious users will increase their chance of alerting the system of their malicious scheme.

**Impact of RepTrap**

The RepTrap algorithm is a targeted attack by malicious users. The end goal is the flip the target business’ review rating or stars. If the business is currently a good business, the malicious users will change the business’ stars to be a bad business (less than 3 stars). RepTrap can also work the other way. Malicious users can flip a business that the crowd has reviewed as a bad business to a good business (greater or equal to 3 stars). Figures 1 and 3 show the required resources to trap the top %10 and top %1.6 business in the yelp data. With a group of people or fake accounts ranging from 34 to 180, nearly every business in the yelp dataset can be target and flipped. The impact that this small set of group has can change public data that all users on the system will see and use in their decision making on products.

**Conclusion**

Crowd sourced data can provide search engines and companies easy access to large amounts of data. The work of creating and updating data is spread out over all the users who use the system. The users are then able to access the knowledge and opinions of others to help their own decisions. The general use of majority voting can be improved for security by assigning user trust to the users of the system. This allows more trust to be given to users who have a good record of voting correctly with the rest of the crowd. It also hinders malicious users by limiting the amount of trust in those with a small number of votes. User trust can still be attacked. The RepTrap algorithm has been shown to be a great improvement in required resources by the attackers. This algorithm of voting can also be improved by skipping the two trap stages and directly voting honestly on items to increase the total malicious trust.

**References**

Q. Feng, Y. Sun, L. Liu, Y. Yang, Y. Dai, "Voting Systems with Trust Mechanisms in Cyberspace: Vulnerabilities and Defenses", 2010.