Privacy-Preserving Polling using Playing Cards

Sid Stamm and Markus Jakobsson
{sstamm, markus}@indiana.edu

Indiana University

Abstract. Visualizing protocols is not only useful as a step towards understanding and ensuring security properties, but is also a beneficial tool to communicate notions of security to decision makers and technical people outside the field of cryptography. We present a simple card game that is a visualization for a secure protocol for private polling where it is simple to see that individual responses cannot be traced back to a respondent, and cheating is irrational. We use visualization tricks to illustrate a somewhat complex protocol, namely the Cryptographic Randomized Response Technique protocol of Lipmaa et al. While our tools — commitments and cut-and-choose — are well known, our construction for oblivious transfer using playing cards is new. As part of visualizing the protocol, we have been able to show that, while cut-and-choose protocols normally get more secure with an increasing number of choices, the protocol we consider — surprisingly — does not. This is true for our visualization of the protocol and for the *real* protocol.

Keywords: card game, polls, privacy, randomized response technique, rational equilibrium, voting.

1 Introduction

Almost without exception, when a new type of cryptographic protocol is designed, it is done in the absence of firm definitions. In some cases, such as mix networks, the complexity of the problem is an obstruction to good definitions; in other cases, such as for many wireless security protocols, the protocol designers have more network knowledge than background in theoretical cryptography, and so, definitions become secondary. In the absence of good definitions, a visualization of the security properties is vital as a sanity check; this corresponds to a visualization of the protocol as well.

Visualizing protocols is a beneficial tool to explain the inner workings of secure systems to those who are not professional cryptographers [1–3]. Therefore, we believe that developing simple and meaningful visual

metaphors for security protocols is an important line of work. In this paper, we use visualization tricks to illustrate a somewhat complex protocol, namely the Cryptographic Randomized Response Technique protocol of Lipmaa et al [4]. Our tools (commitments, oblivious transfer and cut-and-choose) are well known in the cryptography arena, but our construction of oblivious transfer using playing cards is new. As part of visualizing the protocol, we have been able to show what the best parameter choices are.

2 Setting and Problem

You sense a little workplace tension among your employees and want to know if it is a result if your management skills. To figure this out, you want to ask all your employees if you are a good boss. A poll like this may not work as desired. How many employees will be willing to say directly to your face that you are not a good boss? You need to find a polling technique that will preserve their privacy and also give you an accurate account of workplace satisfaction.

A poll is a useful tool used to find out something about a population but may not work as intended when the respondents are asked stigmatizing questions. They may feel inclined to lie when they do not want anybody to know their responses. The individual's concern for their privacy is important and we need a technique to keep their responses private while also maintaining an accurate poll.

We present a simple card game that will maintain this privacy. In our card game, all respondents will follow a set of rules to make sure the poll is successful in maintaining the privacy of the respondents as well as the integrity of our results. In order to ensure this, we also present a mechanism to catch cheaters—those who may not wish to follow the rules in order to alter the results of the poll—and in fact show that cheating is irrational in our game. A version of the same protocol using modular arithmetic was demonstrated by Lipmaa et al [4] in their 2003 publication *Cryptographic Randomized Response Technique.*

3 The Game

We're going to show how a respondent can provide three cards chosen to represent a probability instead of a vote thus removing my ability to be certain how the specific respondent voted. To begin, let's assume a poll is being taken, and the interviewer is called Ira. He asks a question of

Rachel, who is the respondent, "Am I a good boss?" This is a stigmatizing question since most respondents will refuse to say "No" in fear of being fired.

To hide the actual respondent's choice, Ira asks Rachel to provide him with three cards: two representing her honest answer and one that is a lie. We will assume for the sake of this example that red cards mean NO and black cards mean YES. Rachel wants to respond NO so she provides Ira with two red cards representing the truth and a black one for the lie (Figure 1).

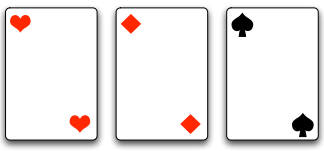

Fig. 1. A sample three-card response representing "NO" (Red)

If she gives the cards to Ira face-down, and he picks one at random, then he doesn't know for sure how she wanted to vote. Rachel cannot change the cards once she presents them to Ira face-down — she commits to them (much like commitment [5]). Ira uses this cut-and-choose method of selecting a card at random to pick a card. In this way, Rachel's privacy is preserved—she can vote either yes or no and Ira still has a chance of revealing a black or red card. In this situation Rachel's honest vote has a $\frac{2}{3}$ chance of being counted the way she intended.

How would Ira sift through these votes where some were recorded as lies to figure out how many people *really* meant to vote YES? He knows that $\frac{1}{3}$ of the votes for YES really should have been NO and vice versa, since a third of the cards he selected were probably the "lie" cards the respondents were asked to provide.

Lets say that Ira takes a poll of 100 people. If 70 of them *mean* YES, then that will result in $70 \times \frac{2}{3} = 46.7$ *measured* yes-votes. The other third (23.3) of the yes-votes will be measured as NO. The remaining 30 voters meant NO, so $30 \times \frac{1}{3} = 10$ of them will be measured YES and the rest (20) will be measured as NO. Thus, Ira measures $46.7 + 10 = 56.7$ yes-votes and $20 + 23.3 = 43.3$ no-votes. What does this mean? If Ira takes a poll measuring 56.7 YES and 43.3 NO, he can assume that 70 out of the 100 people answered YES.

This works in both directions. If he measures 65 YES and 35 NO, this means that the actual yes and no votes (y, n) can be represented in this fashion:

$$\begin{aligned} 65 &= \tfrac{2}{3}y + \tfrac{1}{3}n \\ 35 &= \tfrac{2}{3}n + \tfrac{1}{3}y \end{aligned} \tag{1}$$

In this case, solving for y and n yields 95 actual YES votes and 5 actual NO votes. The lies from each side kind of switched sides and Ira was able to estimate the real responses from those. This oblivious transfer technique [6] is used to keep the vote private; Ira does not know which card in each response was a lie, and the one revealed is at random. He's assuming here that everyone followed the rules, though.

3.1 Stuffing the Ballot Box

Perhaps Rachel wants to fool Ira into thinking he is a worse guesser than he really is. She cannot affect how others play the game, but she can *cheat* in her own game by breaking the rules. To increase the chance that her vote is counted the way she wants (to have more affect on the poll) she can cheat and instead provide Ira with all three red cards for her response. Ira will still think he only has a $\frac{2}{3}$ chance of guessing her answer, when in actuality he will always pick a card that represents her vote. This would give her red vote a 100% chance of being counted as red—which is much better for her, but affects the poll in a bad way. We will call this method of cheating "Stuffing the Ballot Box."

For example, if 45% mean to vote NO, 55% mean YES and 25% of the NO-sayers cheat (by providing all red cards), then Ira gets the wrong idea: he records 52.08% votes as NO. This is because he will record one third of the intended YES votes as NO ($55\% \times \frac{1}{3}$), two thirds of the non-cheating NO votes ($45\% \times \frac{3}{4} \times \frac{2}{3}$) and *all* of the cheating NO votes ($45\% \times \frac{1}{4}$). When he tries to recover the percentage that intended NO (using equation 1) it will seem to him that 56.24% voted NO, which is higher than the actual 45%.

If nobody had cheated in this scenario, Ira would have recorded $45\% \times \frac{2}{3} + 55\% \times \frac{1}{3} = 48.33\%$ NO and $55\% \times \frac{2}{3} + 45\% \times \frac{1}{3} = 51.67\%$ YES votes. Then he would have recovered it (using equation 1) and ended up with 55% and 45%. We need to be sure that all of the respondents will follow

the rules, or Ira won't be able to figure out how many people voted each way in the poll.

3.2 How To Stop Ballot Box Stuffing

Instead of just providing the three-card response, Ira now invites Rachel to provide a "truth" card for her response. A *truth card* is used to indicate whether a response is honest or a lie. Lying is different than cheating: cheating changes the results of a poll and we wish to prohibit this behavior. Rachel is allowed to lie in order to help address privacy issues, but she must be honest about whether or not she is lying—this of course makes it less of a lie.

Rachel tells Ira that she is lying by providing a truth card. If the truth card is black, the response is meant to be honest. Otherwise the truth card is red and the response is the *opposite* of what was intended, or a complete lie. For example, if Rachel wants to vote red, she can respond with a majority of red cards along with a black truth card, or she can also respond with a majority of black with a red truth card (Figure 2).

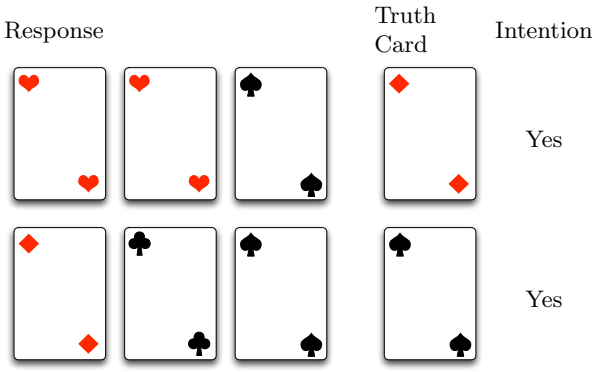

Fig. 2. Two Different but Equivalent Responses

Now when Ira wants to record her response, he must choose one of the three response cards and the truth card to reveal. If the truth card is black ("honest"), then Ira records her vote as the card he revealed. If the truth card is red ("lie") he records the opposite of what he reveals.

By itself, the truth card doesn't change the way the poll works mathematically, but Ira can use the truth card to provide a way of inspecting the responses to see if they follow the rules.

3.3 Enforcing the Rules

Instead of just asking for the minimum number of cards needed to determine a vote, Ira will now ask Rachel for an additional row of response. Figure 3 shows a sample response with all cards turned over to see that the two rows indicate the same vote.

If Rachel answers the question properly, she provides two rows each containing three response cards and one truth card (Figure 3). Ira then chooses one of her two rows at random and reveals one vote card and its corresponding truth card (Figure 4). He then reveals the three vote cards in the second row but *not* the truth card. From this second row he cannot discern Rachel's intended answer since he does not see the corresponding truth card which may be either red or black. In this second row, however, he can see if Rachel is following his rules (two of one color and one of another).

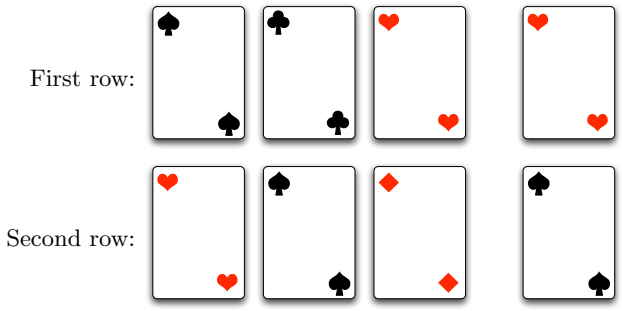

Fig. 3. A single two-row response meaning "NO"

The probability that Ira chooses the right answer from the first row (where he revealed just one vote card and the truth card), however, is still the same ($\frac{4}{6}$ or 66%). In the example shown in Figure 4, Ira would record red as the vote since he revealed a red card and the corresponding truth card was black.

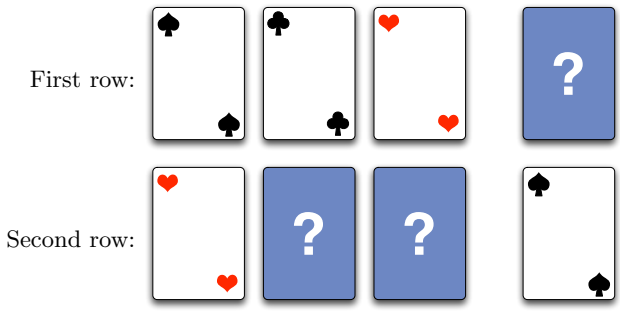

Fig. 4. What Ira sees of Rachel's vote (same cards as Figure 3)

Because of the truth cards, Ira has the ability to look at three more cards without learning anything about Rachel's vote (the first row in Figure 4). Here he is able to reveal the three response cards in the first row, yet since he does not see the corresponding truth card, the only thing he can learn about Rachel by doing this is whether or not she is following our rules. This provides Ira with a mechanism to catch cheaters. He then has the opportunity to disregard votes from anyone who does not follow the rules.

Ira will not always catch the cheaters though, if cheaters only cheat in one of the two rows. Since he only gets to examine one of the two rows for compliance, he can only catch only one out of every two cheaters! Since his row selection is random, only half the time will he pick the "cheat" row to examine. We can increase the chance that he catches cheaters by adding more rows. For example, if Ira asks Rachel for three response rows instead of just two, he can inspect two out of the three rows for conformity (Figure 5). This gives him a 66% chance of catching each cheater since he gets to inspect both the first and third rows. This extends our use of the cut-and-choose technique to rows *and* columns in order to catch cheaters yet still maintain privacy.

Revealing the extra cards in this fashion does not "leak" Rachel's vote to Ira. That is, he does not learn anything about her based on the cards he reveals in the rows where the truth card is not turned over. In no case will he ever see *all* of the response cards in a row *and* the truth card. Of course, the truth card alone means nothing, and without the truth card nothing can be discerned about a row. Also, the truth cards in each row are completely independent: for example, having a black truth card in one

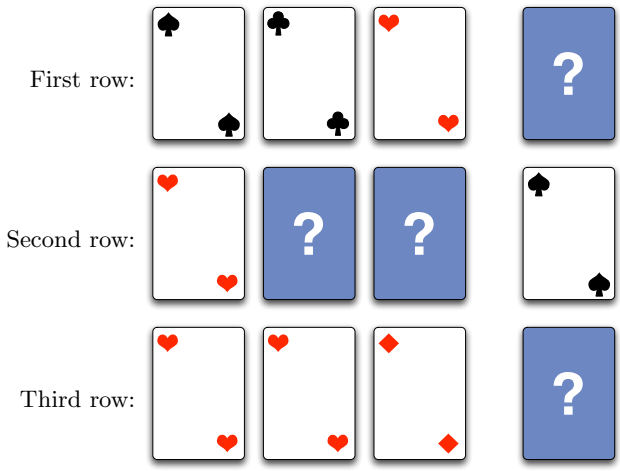

Fig. 5. Ira catches a cheater using a three-row response

row does not effect the probability of having red in another row. Rachel's vote remains private, and if she follows the rules of Ira's game he will be able to conduct a poll in this fashion where cheating is irrational.

4 Cheating is Irrational

Let's say Rachel follows the rules of our game and she provides three rows of response with two honest, one lie and one truth card in each row. The chances that her vote is recorded correctly are $\frac{2}{3}$ or 66%. Her vote will always be counted (one way or another) if she doesn't cheat, so effectively her vote is "worth" 66%.

However if Rachel cheats by breaking the rules in *one* of the three rows, two out of three times she will be caught and her vote won't affect the poll at all. In this case, her vote is only effectively worth 33%, since two out of three times it will be thrown away. The only benefit to cheating is that her vote will definitely be recorded as her desired color. But this benefit still does not outweigh the good chance she will be caught.

That explains how cheating is irrational when Rachel cheats once (in only one of the three rows), but what if she broke the rules in two of the rows? Ira will *always* see one of the cheat rows, since two of the three are always turned up. If she cheated in more than two rows, she is guaranteed

to be caught! An educated cheater would not even think about attempting to cheat in more than one row.

4.1 Changing the Rules

What if we change the game so Rachel only gives Ira two response rows (example in Figure 6)? This means that 50% of the time, Ira will catch the cheat and 50% he will not: she has a better chance of her vote counting. With this 50% chance of her vote affecting the poll, and a 100% chance of it affecting the poll the way she wants, her cheat still isn't "worth" more than an honest vote would have been (66%). So two rows would have been enough to prevent cheating. We chose to use three rows since it is much easier to catch cheaters and throw out the votes: their vote has far less effect on the poll since it is easier to catch.

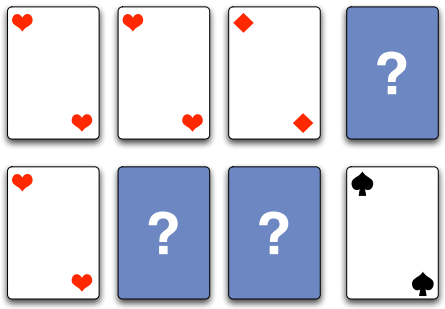

Fig. 6. Rachel Cheated

In a two-row case (Figure 6, the cheater still has a 50% chance of being caught, and so their effective vote still won't be as good as an honest vote worth 66%.

If we add more rows to Rachel's response, Ira's chance of catching her cheating increase, so it would seem that the only way to make cheating any more worthwhile would be to reduce the number of rows in the response to one. In this case, Ira is back to the three-card game without the truth cards or cheat detection, so he would not want to conduct a poll in this fashion.

The only two variables left that might change the worth of a cheater's vote are the number of cards in each response row and the ratio of honest cards to lie cards. Up until now we've explained what happens when three

cards are used with the proportion of honest cards being $\frac{2}{3}$. Adding more cards just provides more precision on the ratio, so we could effectively maintain the same poll with 900-card rows where 600 are honest and 300 are lies.

Changing the response cards to have a lower proportion of honest cards (closer to half honest and half lies) makes a cheater's vote "worth" more, but it also increases the amount of random noise in the poll. At worst, Ira can ask for half honest and half lie cards, but then he won't be able to recover ANY information about any votes since both the YES and NO votes will contain half red and half black cards.

On the other hand, making the cards have a higher proportion of honest cards (closer to 100%) will reduce the effect of cheating by making it more likely that Ira will pick the right color card. At the same time, the rules are well known to Ira, so if the proportion is close to 100%, he can start to make assumptions about the cards he reveals.

The most value we can give to a cheater's response is to make it "worth" the same as a legitimate vote: it can never count for more unless we throw out the rules! In order to make a cheat worth the same as an honest vote, we would have to make the response be half one color and half the other—but then the poll would be useless! All votes would be counted the same (except for the cheaters), and thus the card game would become a game where only cheaters votes mattered. Who would conduct a poll like that?

We've chosen two-thirds as the proportion of honest cards because it is far enough from 100% to maintain privacy while also far enough away from 50% so that the poll still reports useful population sizes.

5 Conclusion

We have shown a game that can be used to conduct polls where the privacy of the respondents (Rachel) is preserved; in other words the interviewer (Ira) learns nothing about the individual respondents' responses, but can later recover the population sizes of how many people voted one way or the other [4]. We have also developed the game such that it is rational for the respondents to follow the rules of the game — thus providing better accuracy than one where cheaters may be motivated to try affecting the results in their favor by cheating.

Using classical techniques, we have constructed a privacy-preserving poll using a card game to visualize our protocol. Respondents in our game cast their votes with a specified number of commitments to cards using

a predefined rule set. This, combined with the cut-and-choose techniques provide privacy to the respondent while also enabling the interviewer to catch cheaters. While even though cut-and-choose usually provides more security with more choices (or cards), that is not the case with our protocol; cheating is irrational with any number of rows in our card game.

References

1. Quisquater, J.J., Quisquater, M., Quisquater, M., Quisquater, M., Guillou, L.C., Guillou, M.A., Guillou, G., Guillou, A., Guillou, G., Guillou, S., Berson, T.A.: How to explain zero-knowledge protocols to your children. In: CRYPTO '89: Proceedings of the 9th Annual International Cryptology Conference on Advances in Cryptology, Springer-Verlag (1990) 628–631
2. Matsumoto, T.: Human-computer cryptography: an attempt. In: CCS '96: Proceedings of the 3rd ACM conference on Computer and communications security, ACM Press (1996) 68–75
3. Shifroni, E., Ginat, D.: Simulation game for teaching communications protocols. In: SIGCSE '97: Proceedings of the twenty-eighth SIGCSE technical symposium on Computer science education, ACM Press (1997) 184–188
4. Ambainis, A., Jakobsson, M., Lipmaa, H.: Cryptographic randomized response techniques. In: Public Key Cryptography. Volume 2947. (2004) 425–438
5. Blum, M.: Coin flipping by telephone a protocol for solving impossible problems. SIGACT News **15** (1983) 23–27
6. Crepeau, C., Kilian, J.: Achieving oblivious transfer using weakened security assumptions. In: IEEE Symposium on Foundations of Computer Science. (1988) 42–52