Deception Behavior Analysis using Mouse Motion Tracking

Agrim Gupta - Y8040

SE367 Introduction to Cognitive Science - Term Project

Mentor: **Prof. Amitabha Mukerjee**

Dept. of Computer Science & Engineering

November 17, 2011

Abstract

Deception is a very typical aspect of human behavior. People exhibit very specific physical, mental and emotional activities when they are telling a lie. Common traits are hesitation, stammering, stalling, looking in the void etc. However, body motions show as much signs of deception as any other trait. Using mouse tracking, we are trying to understand the differences in hand motions when people speak a truth or the lie. Compared to previous studies and and proposed hypothesis, the results do bear significant correlation, and we can safely conclude that people show indirect movements and deviations from the direct paths of body motions when they tell a lie as compared to when they are saying the truth.

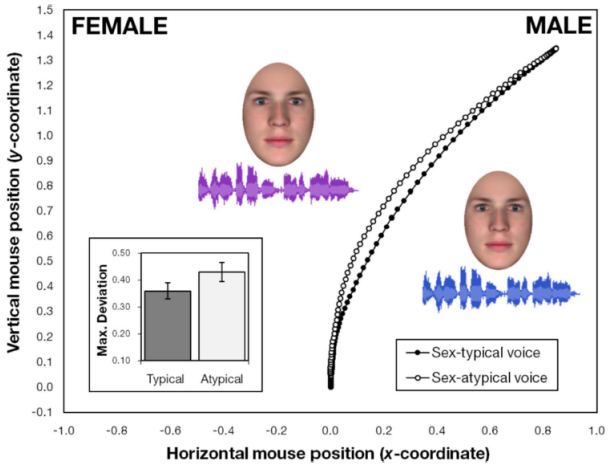
1 Introduction

Traditionally, lie-detection has been a process of cause-and-effect. A question is asked, the person replies, and his mental parameters are noted and ascertained to be true, or false. But not much work has been done in trying to understand the processes that happen during the time the person listens to the question and processes it in his mind.

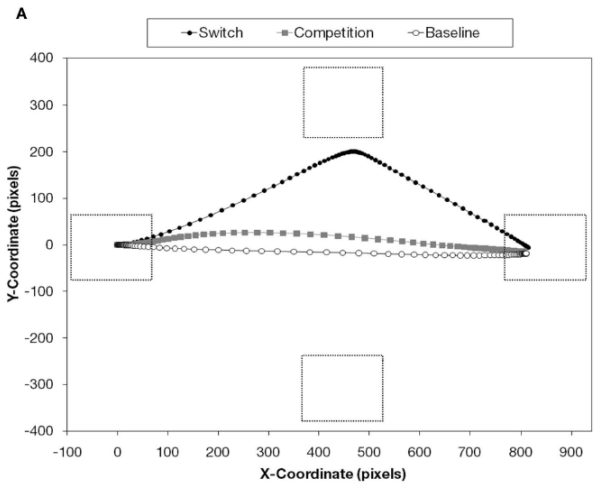
It has been discovered [3] that there exists a deep, intricate and distinctive relationship between the hand-movements of a person and the thought processes that run through his mind when he is saying the truth or a lie. This motion study was done earlier using the game console Nintendo Wii©and its accompanying motion sensitive controller - Wii-mote ©. People are also thinking of applying body motion based research into practice through Xbox360's motion tracking system Kinect ©. These are fairly sophisticated methods to track and understand body motion dynamics - they amount of processing required is very high.

As a result, it is proposed that a new way of tracking deception dynamics behavior using mouse movements shall be studied in this project. Mouse tracking is an efficient and economic way. The equipment required is very commonly available, doesn't require the participant to be aware of the fact that something is being tracked, and the amount of data processing required is under considerable limits and not too sophisticated or computationally intense. Mouse movements have been used earlier

to gauge the typicality of certain images[1] and to verify the continuous nature of language processing in our minds[4].



(a) Mouse tracks illustrating the effect of presence of a competitive option. Participants were shown a picture of an almost sex-atypical face (slightly male deviated) and were asked to confirm their thoughts on the sex of the person in the photo by listening to an additional voice clue. When the the male-typical voice was heard, people readily converged their mouse movements on the 'male' option but when the sex-atypical voice was played, the choice was difficult, and their trajectories showed significant and obvious deviations.



(b) Mouse tracking data showing the effect of the presence of a distracter as a competitor. The participants were told to move the mouse from the left box to a green box. Distractions include like suddenly changing the box colors, inter-changing the positions of the boxes, and also adding multiple boxes of the same color. It was shown that different kinds of distractions produce different grades of responses. Ranging from mild stalling to complete deviation.

Figure 1: Different studies performed for understanding motion dynamics as a consequence of reaction and thought-flow Image source Freeman[6]

2 Problem Formulation

2.1 Idea

A simple response-time curve is not enough for predicting the truth-lie behavior if 'hesitation' is what we are trying to analyze. Neither is an accumulated response sheet of the responses coupled with comparison from a calibrated response of an EEG monitor - as the participants / interrogatees might show unnecessary tension and nervousness traits due to discomfort from the devices and sensors attached to the body. A mouse-movement experiment provides high-fidelity experimental data which captures the thought process in-terms of the mouse-path and is very easily acquirable owing to the ease of operation of the mouse amongst the general public.

So the main concern and objective transforms to understanding how people move a mouse when choosing an answer on the screen that is a truth or a lie. It is expected that people will show distinct mouse behavior when choosing an answer on-screen. A mouse was used to make the task simple and take the oddness out of it. The task is non-intrusive and the participant may not even feel that the tool which he is using so comfortably, is the tool that is understanding him the most.

2.2 Methodology

The core task was simple - to make a person read and/or listen to a yes/no question, and ask the person to answer it either truthfully, or lie about it. So this makes 4 possible nature of answers:

1. True/Yes
2. False/Yes
3. True/No
4. False/No

The participant will be asked to move the mouse pointer towards the answer as soon he/she starts hearing/reading the question. This is because language is evaluated continuously as we keep on assimilating it and the response of the participant will be dependant on the portion of the question heard or read, and the what the participant predicts it to be. The questions asked in this process will be very obvious and rudimentary and will have very easily verifiable factual answers. The list of questions used for the experiment can be seen in the Appendix A.

The experiment is designed in a software called Mouse tracker by J.B. Freeman [5]. The main constructs are specifying the locations of the stimulus, starting point of the mouse pointer, numbers and positions of the responses. This can be seen in the screenshot presented in Figure 2.

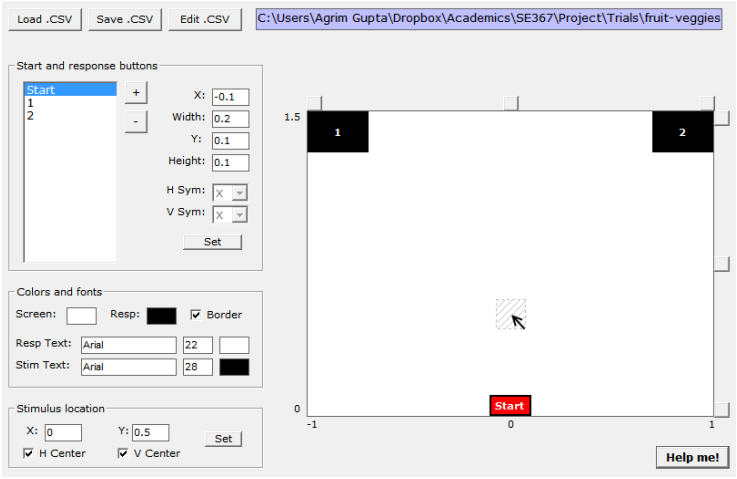


Figure 2: A screen-shot depicting the mouse-tracker designer software for designing the experiments

As found out by Dale[1;2], there is a very distinct difference in the patterns of all four choice options. e.g. The True/Yes option has the most direct path taken from the source to the correct choice. While the False/Yes shows the highest deviation/competition between the two choices. Dale analyzed the responses in terms of the deviations from straight paths and deviations when compared amongst themselves. This can be more clearly understood from Figure 3

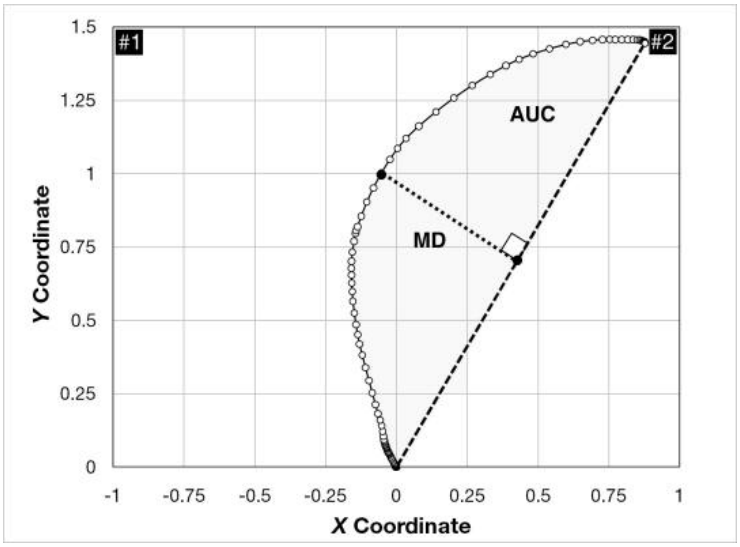


Figure 3: Quantitative analysis of the deviation

The mouse tracking is initially processed in the analyzer part of MouseTracker©. We get the mouse trajectories are x and y coordinates on a normalized time-scale

basis. This helps us much in comparing results and plotting time-based results. A sample of the process of analyzing is shown as a screenshot in Figure 4

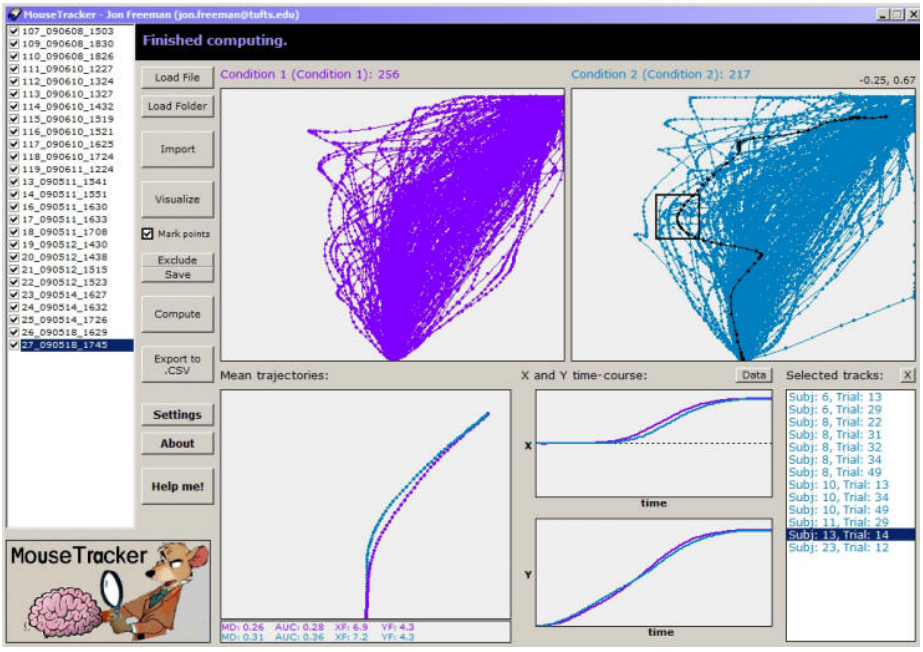


Figure 4: Mouse tracker analyzer for analyzing and visualizing the mouse tracking data

3 Work, Survey & Results

3.1 Survey

Two sets of 15 questions each were prepared and the participants were asked to answer questions in one set truthfully, and were told to lie to the questions in the other set. The participant was free to choose which set he wants to answer to first, and which set he wants to answer truthfully/fakely. This step was taken to ensure that the response data is substantially random in nature and not much bias is present in the data. The questions can be found in Appendix A. The total number of participants were 15. As such, the number or participants is irrelevant, we are much interested in individual results.

3.2 Work

Further data processing was done in Matlab. Several ideas were experimented and implemented and a few survived. The following sections will illustrate the work done and the results obtained. Some of the ideas were the following:

- Understanding time-based motion of the mouse pointer on the x-axis as well as the y-axis, separately. The time-based motion tells us about the amount of

decision time taken up and also the key factor about the presence of hesitation - which is the first indicator of a lie - necessary, but not sufficient. This can be done by making plots like in Figure 5.

- Comparing the average behavior of the trajectories in the Truth and Lie cases (See 6). This gives us a clear comparison into the differences in the dynamics involved in truth and lie processes.
- Comparing complete x v/s y trajectories in the case of Truth and Lie. This is the main curve that we are interested in. It provides a direct visual estimate of the deviations in each tracking trial.
- For the final step we can try and implement a simple machine learning concept to build a calibration curve for a participant and then predict whether the next answer by the participant is a truth or a lie. How we do this, we shall discuss in the next section

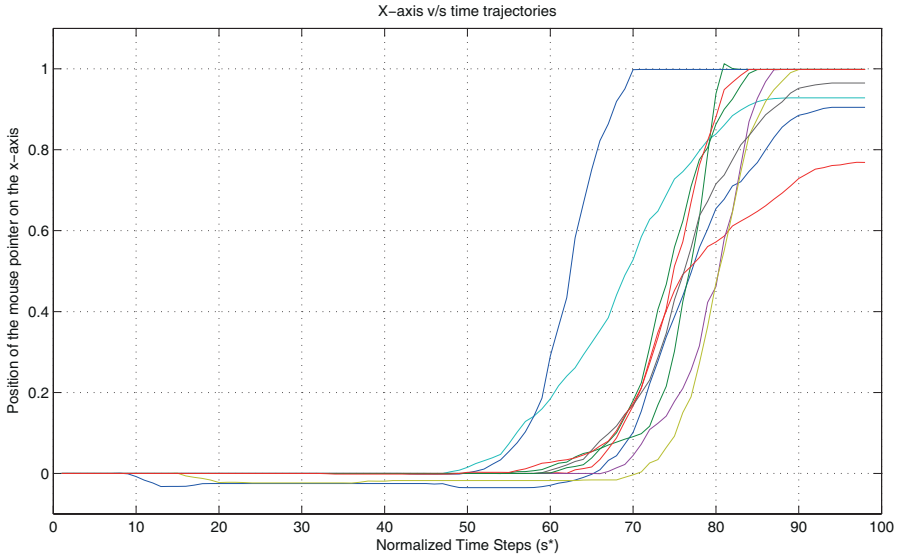


Figure 5: X-Lie curve. X-axis components of the mouse movement trajectories for one participant while lying. The time-scale is normalized. We see that people keep their mouse-pointers almost in the center of the screen (where they begin) for a significant amount of time while they decide the option to choose. Once decided, they move the mouse fairly quickly until they reach somewhere near the response they have chosen.

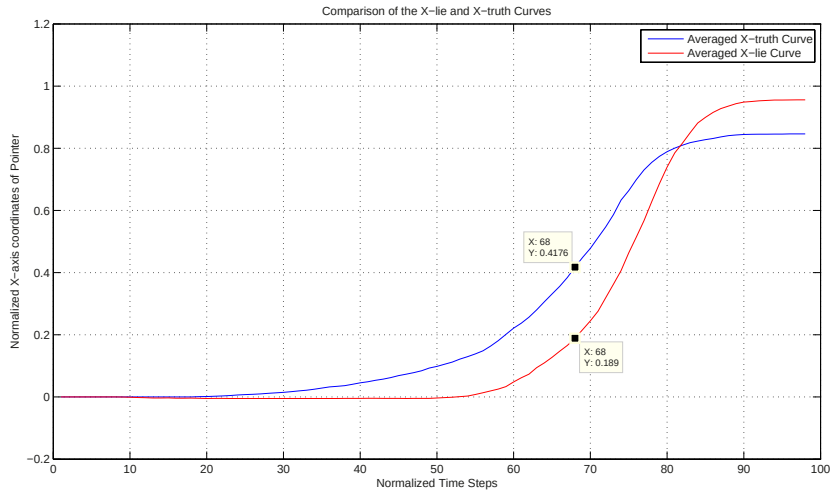


Figure 6: Comparison of the X-lie and X-truth curve. Note that when saying the truth, people start moving their mouse much more earlier as compared to when they are lying - they require a longer time for negating true answer in their mind, deciding to go for the lie and finally moving the mouse. This additional processes in between are what we set out to detect in the beginning and we have conclusive evidence that there is a significant bit of hesitation in people's body movements when lying.

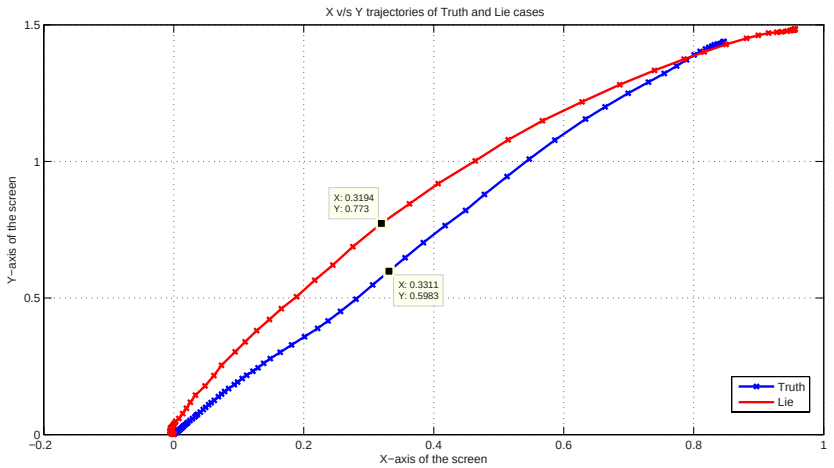


Figure 7: Comparison of the X/Y-lie and X/Y-truth curve. The lie curve deviates towards the competitive true answer before converging onto the lie that the participants had to choose. This indicates a tussle in the mind over choosing from an option that is the answer and an option that should be answered.

3.3 Calculations

To determine whether a response is true or a lie, we need to first know the pattern in which a participant answers questions. This is a simple process of machine learning. Here we are considering the average of the coordinates of the trajectories of some of the responses given by the participant and evaluating the correctness of the rest of the responses and checking for the validity of our hypothesis.

The procedure that we employ to check for each response is the following:

1. Evaluate the average trajectory for the first n true and lie responses each
2. Pick up the trajectory of the next response
3. Find out the distance between the vectors of the average trajectories (both truth and lie), and the trajectory we want to evaluate
4. If the distance of the current response is closer to the truth trajectory, we say that the current response is true, else, if the distance of the current trajectory is closer to the lie trajectory, we say that the person has lied

Based on this simple idea, we evaluated the responses of the participants and in each case the following parameters were noted down:

- Number of truth responses being evaluated as truths
- Number of truth responses being evaluated as lies (false negatives)
- Number of lies being evaluated as lies
- Number of lies being evaluated as truths (false positives)

3.4 Results

Based on the above described procedures, we found out the following results:

- The accuracy of our algorithm varies with variation in the amount of initial responses provided to build the calibration curve
- This variation is somewhat random. Although increasing the initial sample size increases the accuracy, the pattern is not very reliable and some amount of judgement is required when selecting the sample size
- The accuracy to detect correct truths and correct lies is nearly always more than 65% and in some cases, the accuracy can even go up to as high as 90%. These results are being considered with a sample size that is $(1/5)^{th}$ the size of the total number of responses.
- If the sample size is increased to 50% of the responses, the baseline accuracy improves to 75%.
- Different vector distance finding algorithms lead to different levels of accuracy. It was found that the 'chebyshev' method in Matlab was giving the best results compared to other distance methods

4 Conclusions

The results as discussed above show very good promise and indicate a direct evidence of the validity of our hypothesis. It was found that **'Lie detection'** can be done with accuracies more than 75% using a simple mouse motion tracking method. This level of correlation is very good considering that the tests were not even performed in ideal test conditions. If the way the test is conducted is improved further, we can see a much higher ratio of accurate lie detections. This is a totally non-intrusive and highly economical method to perform deception analysis and other cognitive studies that require understanding the motion dynamics part of a cognitive process.

The main drawbacks with the experiment were mostly due to the limitation of the features and flexibility of the software used. It was very difficult to fully randomize the position, and nature of the responses. Correct responses always remained in one position, and incorrect responses in the other. Thus, towards the end of the experiment, the participant might become aware of this fact and this might introduce a bias of deliberation into the results. However, even after so many hurdles, the results are coming out more than satisfactory - which is a good sign.

References

[1] R. Dale, C. Kehoe, and M.J. Spivey. Graded motor responses in the time course of categorizing atypical exemplars. *Memory & Cognition*, 35(1):15–28, 2007.

[2] R. Dale, J. Roche, K. Snyder, and R. McCall. Exploring action dynamics as an index of paired-associate learning. *PLoS One*, 3(3):e1728, 2008.

[3] N.D. Duran, R. Dale, and D.S. McNamara. The action dynamics of overcoming the truth. *Psychonomic bulletin & review*, 17(4):486–491, 2010.

[4] T.A. Farmer, S.A. Cargill, N.C. Hindy, R. Dale, and M.J. Spivey. Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science*, 31(5):889–909, 2007.

[5] J.B. Freeman. Mousetracker software. `http://mousetracker.jbfreeman.net/`, October 2011.

[6] J.B. Freeman, R. Dale, and T.A. Farmer. Hand in motion reveals mind in motion. *Frontiers in Psychology*, 2, 2011.

A Question List

1. Rabbits never turn down a carrot
2. It is night right now
3. You are an IIT-ian
4. Are you still there?
5. Do you watch tv
6. Newton discovered laws of motion

7. Smoking is injurious for health
8. Practice makes a man perfect
9. Do you study at IIT Kanpur?
10. Do you watch movies?
11. Do you drink coffee?
12. We get wool from sheep?
13. Do you like treats?
14. Is it winter?
15. The earth is round?
16. The sun is 5 billion years old
17. Ac's give cold air?
18. Is there a restaurant around?
19. You have a mobile phone?
20. Doors have locks?
21. We travel by cycle
22. Sachin is a great cricketer
23. Formula of water is h2o
24. India hosted an f1 race for the first time
25. Are you from India?
26. Is it winter?
27. The earth is round?
28. The universe is 13 billion years old