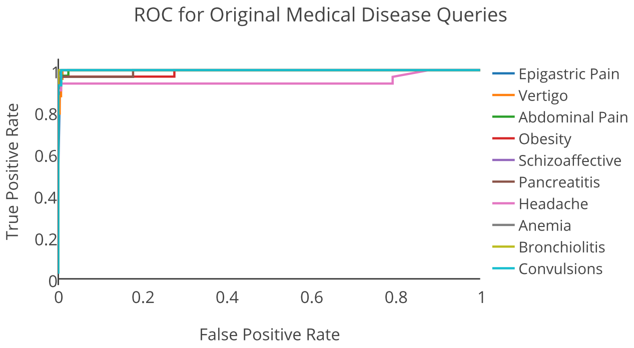

Figure 1: (Pre-obfuscation) Rate of true positives vs. false positives for classifying memory request traces of medical disease queries. Each query is for the average cost of different admitting patient diagnoses. This shows our modeled attacker's classification has near perfect accuracy when observing the original queries execution.

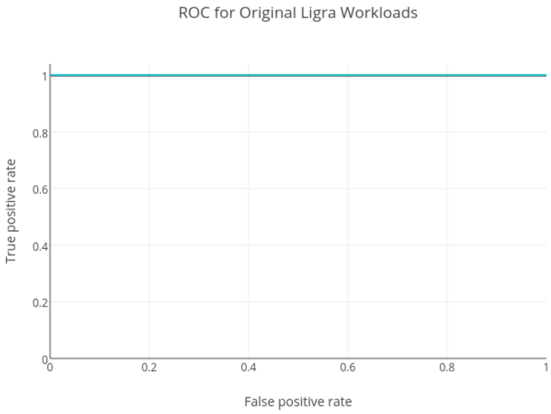

Figure 3: (Pre-obfuscation) Rate of true positives vs. false positives for classifying memory request of different graph applications and inputs. The attacker's classification is perfect.

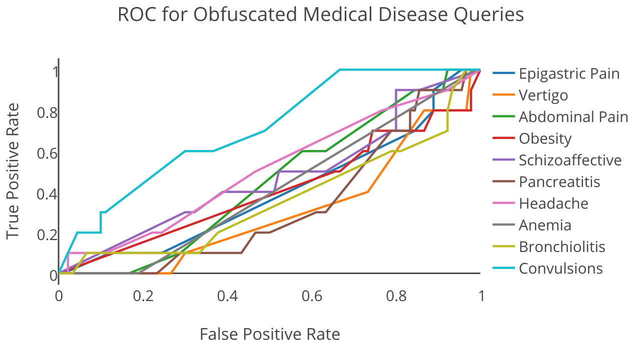

Figure 2: (Post-obfuscation) Rate of true positives vs. false positives for classifying obfuscated memory request traces of the same medical disease queries using a privacy parameter of $\epsilon = 0.1$. This shows our modeled attacker's classification is much more difficult after applying our obfuscation scheme.

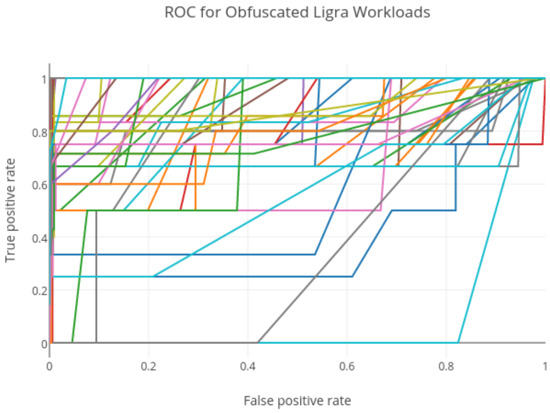

Figure 4: (Post-obfuscation) Rate of true positives vs. false positives for classifying obfuscated memory request traces of the same graph applications and inputs using a privacy parameter of $\epsilon = 1$. This shows our modeled attacker's misclassification rate has significantly increased.

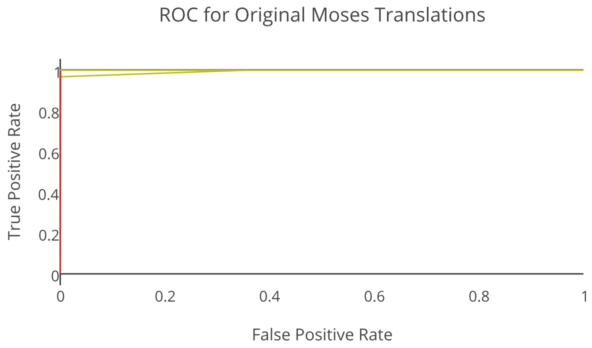

Figure 5: (Pre-obfuscation) Rate of true positives vs. false positives for classifying memory request traces of 50 sentence translations. This shows our modeled attacker's classification has near perfect accuracy when observing the original queries execution.

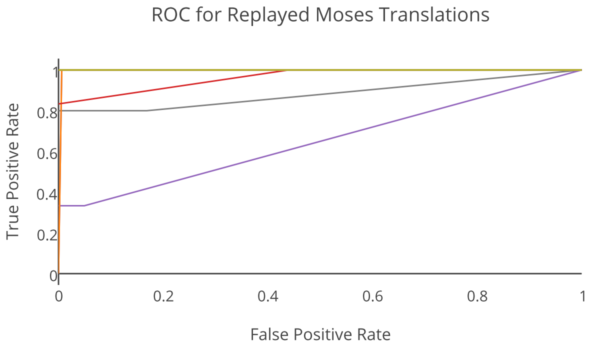

Figure 6: (Post-replay) Rate of true positives vs. false positives for classifying obfuscated memory request traces of the same translations using our replay model. This shows our modeled attacker's classification is not affected significantly by the replay mechanism itself, as many of the inputs are still perfectly predicted. Any privacy we obtain is from the noisy trace computed.

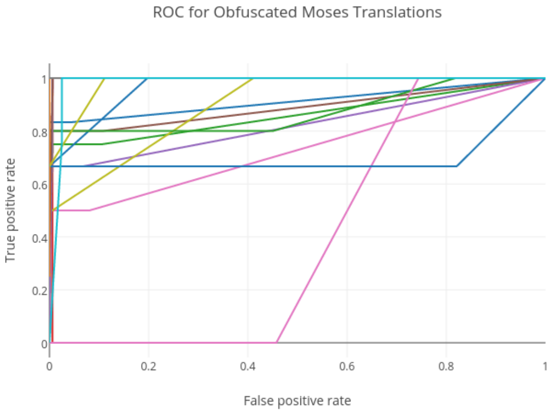

Figure 7: (Post-obfuscation) Rate of true positives vs. false positives for classifying obfuscated memory request traces of the same translations using a privacy parameter of $\epsilon = 10$. This shows that even at our lowest privacy guarantee, our modeled attacker's classification is affected significantly and the misclassification rate increases.