An online attack against Wiesner's quantum money

Andrew Lutomirski
*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139**
(Dated: September 31, 2010)

Wiesner's quantum money [5] is a simple, information-theoretically secure quantum cryptographic protocol. In his protocol, a mint issues quantum bills and anyone can query the mint to authenticate a bill. If the mint returns bogus bills when it is asked to authenticate them, then the protocol can be broken in linear time.

INTRODUCTION

In [5], Wiesner proposed a protocol for information-theoretically secure private-key quantum money. A mint can choose a security parameter n and generate a random n-qubit state. (Each qubit is independently and uniformly drawn from $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$.) The mint assigns that state a unique serial number and declares it to be a \$20 quantum bill. (The \$20 is arbitrary.) To verify a quantum bill, a merchant sends the bill to back to the mint. The mint looks up the classical description of the state matching the serial number of the bill and projects the quantum state being tested onto that state. A "**VALID**" result (the state being tested matched the description) means that the bill is valid and an "**INVALID**" result means that the bill was counterfeit or damaged. (The mint needs to maintain a secret database of the description of the random state corresponding to each serial number.)

This protocol is information-theoretically secure. The no-cloning theorem implies that an attacker cannot perfectly copy a quantum bill, and the bounds in [1] mean that the probability that an approximate copy appears valid drop exponentially as a function of n.

There are many recent papers based on the idea of attacking otherwise secure *classical* cryptographic protocols by various side channels or online attacks. For example, in 2002, Vaudenay showed that a commonly used form of symmetric cipher (CBC mode encryption) can be attacked with a small number of queries to an oracle that distinguishes valid encrypted messages from invalid messages with a certain type of error [4]. Rizzo and Duong dramatically showed that these attacks worked against carelessly designed websites and that many current websites are vulnerable [3].

Inspired by Rizzo and Duong's result, I show that, even if the mint has a perfect quantum computer, Wiesner's quantum money is vulnerable to an online attack. If, when asked to verify any bill, the mint returns the bill *even if that bill was invalid*, then a small number of queries to the mint can be used to copy a bill.

THE ATTACK

For Wiesner's quantum money to be useful, the mint must offer a service that anyone can use to verify quantum bills. Morris (presumably a merchant) sends the mint a quantum bill. The mint either answers **VALID** and returns the bill to Morris or answers **INVALID**. In the **INVALID** case, if the mint destroys the counterfeit bill, then all is well. If, on the other hand, the mint returns the counterfeit bill to Morris, then the entire protocol can be broken in linear time.

We can formalize the quantum bill as a classical-quantum state $(s, |\$_s\rangle)$ where s is some unique classical serial number and $|\$_s\rangle$ is the random product state chosen by the mint that corresponds to the serial number s. We can write

$$|\$_s\rangle = |\psi_1\rangle|\psi_2\rangle \dots |\psi_n\rangle,$$

where each $|\psi_i\rangle$ depends on s and is drawn from $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$. This means that each $|\psi_i\rangle$ is an eigenstate of either X or Z. (Morris does not know which operator each $|\psi_i\rangle$ is an eigenstate of.) We assume that Morris can send the mint any c-q state $(s, |\phi\rangle)$ and the mint will measure the projector $P_s = |\$_s\rangle\langle\$_s|$. If the outcome is 1, the mint returns $(\mathbf{VALID}, s, P_s|\phi\rangle)$ and if the outcome is 0, the mint returns $(\mathbf{INVALID}, s, (1 - P_s)|\phi\rangle)$ (up to normalization).

If Carla the counterfeiter has a single quantum bill $(s, |\$_s\rangle)$ and can query the mint, then she can break the protocol by learning the state $|\$_s\rangle$ one qubit at a time. To learn the ith qubit, she sends the mint the state $(s, X_i|\$_s\rangle)$. If the mint answers **INVALID**, then the state $|\psi_i\rangle$ was either $|0\rangle$ or $|1\rangle$ (as the other possibilities $|+\rangle$ and $|-\rangle$ are eigenstates of X_i). In this case, the returned state is

$$|\psi_1\rangle \cdots |\psi_{i-1}\rangle|\psi_i^{\perp}\rangle|\psi_{i+1}\rangle \cdots |\psi_n\rangle.$$

But now Carla knows that $|\psi_i^{\perp}\rangle$ is an eigenstate of Z. She applies X_i to recover $|\$_s\rangle$ and measures $|\psi_i\rangle$ in the Z basis to learn whether it is $|0\rangle$ or $|1\rangle$.

If, on the other hand, the mint answers **VALID**, then the state $|\psi_i\rangle$ was either $|-\rangle$ or $|+\rangle$. In this case the mint returns the (undamaged) state $|\$_s\rangle$ to Carla. But now Carla knows that $|\psi_i\rangle$ is an eigenstate of X and she can measure it to learn whether it is $|+\rangle$ or $|-\rangle$.

If Carla repeats this process for $i = 1, \ldots, n$, she will learn the secret description of $|\$_s\rangle$ in exactly n queries to the mint. Once she has done this, she can make as many counterfeit copies of $|\$_s\rangle$ as she wants.

Carla could also use a more generic algorithm such as quantum state restoration to copy the state directly in $2n$ (expected) queries to the mint or single-copy tomography to learn the state in $O(n)$ queries [2].

CONCLUSION

Anyone who implements classical cryptographic protocols needs to be very careful to avoid introducing flaws that bypass the security that the protocol would have offered if correctly implemented. Quantum cryptography is not magically safer.

ACKNOWLEDGMENTS

I was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

---

* Electronic address: luto@mit.edu

[1] Vladimir Bužek and Mark Hillery. Quantum copying: Beyond the no-cloning theorem. *Physical Review A*, 54(3):1844–1852, 1996.

[2] Edward Farhi, David Gosset, Avinatan Hassidim, Andrew Lutomirski, Daniel Nagaj, and Peter Shor. Quantum state restoration and single-copy tomography. 2009, arXiv:0912.3823v1.

[3] Juliano Rizoo and Thai Duong. Practical Padding Oracle Attacks. In *4th USENIX Workshop on Offensive Technologies*, 2010.

[4] Serge Vaudenay. Security Flaws Induced by CBC Padding–Applications to SSL, IPSEC, WTLS... In *EUROCRYPT 2002*, pages 534–545. Springer, 2002.

[5] Stephen Wiesner. Conjugate coding. *SIGACT News*, 15(1):78–88, 1983.