Extended Abstract - The Quantum Trojan Horse Attack

It is often claimed that quantum key distribution (QKD) promises security against eavesdropping that is guaranteed by the laws of physics. However, the existence of various successful eavesdropping attacks demonstrates that this cannot be the full picture. Typically, when proving the security of a particular QKD implementation, additional assumptions about the protocol are called upon. In particular it is common to assume that Eve (the potential eavesdropper) has no access to the internal workings of Alice's device which prepares the quantum state that she sends to Bob. This is sometimes referred to as the *secure laboratory* assumption. However, this assumption does not always hold in the real world. Whilst Alice may be able to *limit* Eve's interaction with her apparatus, there is always one channel that will remain open, which is the fibre through which Alice sends her signals to Bob.

Suppose that Alice and Bob are trying to communicate by the BB84 protocol. Alice encodes her information in states of the form $\left(|0_Z\rangle + e^{i\theta}|1_Z\rangle\right)/\sqrt{2}$, where $\theta \in \{0, \pi\}$ corresponds to the X basis, and $\theta \in \{\pi/2, 3\pi/2\}$ corresponds to the Y basis. Eve may send her own state through the optical fibre into Alice's system, where it may be pick up some information about Alice's choice of θ. This state is then returned to Eve, who makes some measurement on it. Thus she may learn the secret key without disturbing Bob's qubits, and so without being detected.

Defenses against this attack have been suggested using active components, but these may introduce new vectors of attack for Eve. A passive defense using one-way attenuators has been suggested, but this assumed specifically that Eve would use a coherent state. Until now there has been no *complete* security analysis of this problem where Eve may do anything allowed by the laws of physics, including any choice of input state, and where Alice and Bob make use of passive defenses.

We analyse this system by modeling the effect of the attenuator on Eve's state as a CPTP map on some general photonic state, which is then returned to Eve. For a sufficiently large attenuation, the resulting state is restricted to a subspace of the Hilbert space consisting of only a few photons. We can then bound the off-diagonal terms (which carry the information about θ) by enforcing that the state is positive. By supposing that Eve then makes a measurement to determine how many photons she has before analyzing her state, we can bound the contribution of higher-order terms.

By doing this, we present an absolute security bound for any (non-entanglement-assisted) attack strategy that Eve might employ. This gives a secret key rate for Alice and Bob, expressed in terms of one easily measurable parameter, Eve's average received photon number, μ. This shows that the level of attenuation that was found to be sufficient to protect against coherent state attacks may be too small by up to a factor of 100. This result is crucially important to anyone who wishes to implement any BB84-like QKD communication protocol, since there is no passive way of detecting the Trojan Horse Attack, and failure to defend against it will result in absolute vulnerability of the secret key. This work presents a threshold theorem, with the promise that if Alice implements a certain level of attenuation (99.5% attenuation to get a non-zero key rate, compared with the 0.9% required if Eve used a coherent-state attack), then the communication is secure against any (non-entangled) attack. Work continues on the entangled-state attack, which would constitute an absolute bound on security for this type of side-channel.