TrueCrypt's Plausible Deniability (Hidden Volumes) is Theoretically Useless

5-7 minutes

March 2, 2013

When analyzed with game theory, it turns out that TrueCrypt's plausible deniability feature, which lets you hide a second encrypted volume inside the "outer" or normal volume, is useless.

Okay, that's a bit of an exaggeration, but let me explain.

Consider the following situation: You are a dissident under an oppressive government, and you want to encrypt your plans to overthrow the government. You are pretty sure that you will eventually get caught and your government will force you, by torture, to disclose your password. Suppose you decide to use TrueCrypt, and you're wondering whether or not to have a hidden volume.

The government will face a different choice when they catch you. Knowing that TrueCrypt has a hidden volume feature, they have to decide whether or not to continue torturing you (possibly until you die) after you reveal the normal volume's password (and you WILL reveal the password; if you don't believe me, read the ending of Nineteen Eighty Four).

When we write down all combinations of these decisions, we get something like this:

	No Hidden Volume	Hidden Volume
Stop Torturing	Y: -10, G: 9	Y: 10, G: -10
Keep Torturing	Y: -100, G: 10	Y: -50, G: 10

This table assigns a number to your (Y) "reward" and the government's (G) reward for all of the possible decisions:

- If you don't have a hidden volume and the government decides to keep torturing you, you're screwed. You have no way of stopping the torture even if you wanted to, and you probably get killed. So your reward is -100. The government's reward is 10 because although they may have wasted some time torturing you, they have your plans, can prove your guilt, and arrest your accomplices.

- If you don't have a hidden volume, and the government decides to stop torturing you, then your reward is -10 because the government got your plans, proves your guilt, and arrests your accomplices, but you don't get tortured as much and maybe the government lets you become an informant in exchange for letting you live. The government's reward is 9 here, instead of 10 like before, since they aren't as sure that you don't have a hidden volume.

- If you have a hidden volume, and the government stops torturing you, then you get a reward of 10, since the government has not found your plans. The government gets a reward of -10, since they failed to recognize you as a dissident and risk being overthrown.

- If you have a hidden volume and the government keeps torturing you, you get a reward of -50, instead of -100 like before, since in this case you have the option of surrendering the password to stop the torture and you might get out alive. Again, the government gets a reward of 10 since they found your plans.

It's important to note that the actual number's don't really matter here. All that matters is their ordering (greater than or less than). Using numbers is just a convenient way of expressing relative goodness/badness.

Now here's the interesting part. It is a strictly dominant strategy for you to use a hidden volume, and it's a strictly dominant strategy for the government to keep torturing you.

What is a strictly dominant strategy? It's a strategy that is optimal *no matter what the other player does*. For every move the other player can make, it is best for you to go with the strictly dominant strategy. You might be familiar with this concept from the prisoner's dilemma, where the strictly dominant strategy for both players in that game is to rat each other out.

To see why it's a strictly dominant strategy to use a hidden volume, consider both moves the government can make. If the government stops torturing you, then you get a reward of 10 for having a hidden volume and a reward of -10 for not having one. So if the government stops torturing you, it's best if you have a hidden volume. Likewise if the government keeps torturing you, you can either have a reward of -100 (no hidden) or of -50 (hidden). In both cases, no matter what the government chooses to do, it's best if you have a hidden volume.

It's also a strictly dominant strategy for the government to keep torturing you. If you're not using a hidden volume, they can get a reward of 9 (stop torturing) or 10 (keep torturing), so in that case, it's best to keep torturing. Similarly if you are using a hidden volume, they can get a reward of -10 (stop torturing) or 10 (keep torturing), so it's best to keep torturing in that case too. So no matter if you're using a hidden volume or not, the government gets the highest reward by continuing to torture you.

So if you and the government are both rational and self-interested, then you are going to use a hidden volume, and the government is going to keep torturing you.

So, at least in this situation (which arguably is why the feature exists), TrueCrypt's plausible deniability doesn't give dissidents any advantage. In fact, it could make things worse for the dissident if they don't know about the feature and end up being tortured for a password that doesn't even exist.

In other scenarios the feature can be useful. If the attacker has limited resources (i.e. can only torture you for 30 minutes), or if you are "innocent until proven guilty" under the law, then it can be advantageous to use a hidden volume. Just don't recommend TrueCrypt to your friends in North Korea, or at least make sure they use a hidden volume.