Seed Scanning Malware: Why You Should Never Store a Cryptocurrency Seed Phrase in Plain Text

Author	josh
Date	2023-12-06

Overview

Storing cryptographic keys properly is *hard*. Getting it wrong is pretty easy. And there's one major way you can get it wrong – by storing a copy of your cryptocurrency wallet seed in plain text form. Let's look at an educational demo of the sorts of attacks that can compromise your seed phrase – seed scanning malware.

Seed Scanning Malware Basics

How Does This Attack Work?

When I say storing a seed phrase in *plain text*, what does that actually mean? In computer security terms, that means storing the seed data in an *unencrypted* form. Meaning that anyone that can see the file, can see the actual seed. A plain text seed might be stored in a text file on you desktop for example, or in a word document. Some people have even stored screenshots of their seed or a copy in a cloud storage provider.

You might say, well why does this matter? After all, I'm the only one with access to my computer or my phone. Well, that's not always true. You might be affected by malware – unauthorized malicious code running on your system. And in fact, it's quite easy to write code that can scan for seed phrases using common programming tools!

Seed Scanning Code

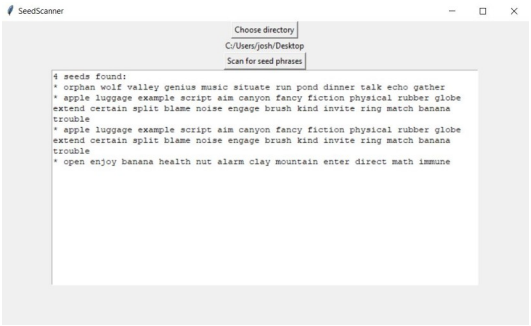
Demonstrating the concept of seed scanning malware is actually rather simple, using software development tools that are common to programmers. In this case, we can use a tool called a *regular expression.* Regular expressions are tools for string *pattern matching*, and are widely used. Developers might use regular expressions to validate a website URL, or a phone number for example. Regular expressions are adept at finding a specified pattern in plain text data, and since we know the pattern of a cryptocurrency *seed phrase*, we can construct a regular expression to search for them in text documents!

A cryptocurrency seed phrase takes the form of 12-24 English words from a 2048 word *dictionary*, specified by the *BIP39* standard. BIP39 wallets are widely implemented throughout the crypto space, and make backing up or importing wallets much easier that older methods of key storage.

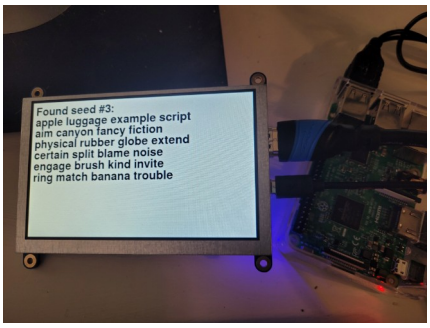
Writing a regular expression to match a cryptocurrency seed phrase is fairly trivial – we search for any instances of the 2048 possible words, and assume that the words will occur in a row (with some whitespace in between). The regular expression is quite large and unwieldy, but actually runs pretty efficiently. We can construct the regexp programatically with the words:

```
with open(self.WORDLIST_FILENAME) as f:
    words = [word.strip() for word in f]

combined_words = "|".join(words)
regex_str = r"((" + combined_words + r")\s*){12,24}"
regex = re.compile(regex_str, flags=re.IGNORECASE|re.MULTILINE)

return regex
```

This Python code constructs the regular expression, which can later be used to search the contents of text files in a particular directory we'd like to search. While this code is quite simple and educational, real seed-scanning malware can be much more sophisticated. Real attackers might employ more in-depth methods for seed searching, and exfiltrating the data into their hands. Further validation could be done to check the validity of the seed beyond the simple pattern, as there's other checks that go into creating a valid seed. There's lots of possibilities for updating this code.



A simple seed auditing graphical demo of the SeedScanner library



A demo of the library and UI running on a Raspberry Pi

Seed Stealing Prevention

Knowing that these sorts of attacks are possible, how does one prevent a seed scanning attack, assuming that our system could be theoretically compromised with malware? Well, key management is complex. But there's some straightforward advice that can apply to this context.

If the seed is for a hot wallet, that is, a wallet created on your phone or PC: in my opinion, it is okay to store a copy of the seed phrase in an *encrypted* form on your PC or phone, such as a secure password manager. You must ensure the encryption key is derived from a long and *strong* passphrase. The reason this is likely okay is that if you're using a hot wallet, the keys are *already* stored by the wallet on that device in an encrypted form. You're simply created an encrypted

backup of those keys that are already stored and generated on a general purpose, networked computing device.

If the seed is for a cold wallet, such as a hardware wallet, you *must* only store a copy of that seed on paper, metal, or some other physical form. Never store a cold wallet seed, even encrypted, on a general-purpose device such as a PC or phone. Doing so breaks the security model of a hardware wallet, whose purpose is to generate and store keys *offline*, away from general purpose computing devices that can run malicious code. If you keep the keys away from a PC or phone, they cannot be scanned for by malware.

Keep Your Keys, Keep Your Coins

With some fun and educational code, we've demonstrated here why it's a bad idea to store a seed phrase in plain text. It only takes the use of common programming tools to construct a demo of malicious code, and the real thing can be much more dangerous. There's a lot of interesting ways to tinker with security and code, but when it comes to your real cryptocurrency keys, you should always be extra, extra careful. Store hot wallet seeds in encrypted form, using strong passphrases and well-vetted software. Keep your cold wallet seeds in physical form only. Proper key management and education means you get to keep your hard earned coins!