

ciphart

memory-*harder* key derivation with easier security interpretation

caveman¹

2021-02-26 20:00:24+00:00

synopsis—*argon2*² is mostly a nice memory-hard key derivation function, as it is simple compared to *scrypt*³, which makes understanding what it is doing easier. but i think that *argon2* is still not nice enough, as i think that its security contribution can be simpler to understand, and i think that it can have *harder* memory-hardness.

currently, if you want to know *argon2*'s contribution to your security against password brute-forcing, you need to survey the industry of application-specific integrated-circuits (asics) in order to find the cost of breaking your password in a given time frame, as done in the *scrypt* paper. the housekeeping of this approach is too expensive as the industry is constantly changing, plus it remains a rough estimate with inadequate theoretical guarantees.

also, *argon2*'s memory-hardness is based on only utilising random-access memories, and not, say, hard-disks. without changing the threat model of the vast majority of users, this is an unnecessary restriction as i show in this paper.

henceforth, i propose *ciphart* — a memory-*harder* key derivation function, with:

- a security contribution that is measured in the unit of shannon's entropy. i.e. when *ciphart* derives a key for you, it tells you that it has *injected* shannon's entropy bits into your derived key, and also tells you its exact quantity. this is possible thanks to my invention "perfect lie" theorem.
- and a memory-*hardness* that comes from the fact that *ciphart* can require crazy amounts of memory, beyond our random-access memory, thanks to it being able to use the hard-disk as well. this is possible thanks to my discovery "cacheable keys".

this is optional, but i extremely like it as it effectively gives me much more security while eventually becoming much faster as well, and the adversary cannot get my cache even if he steals my hard-disks.

libciphart⁴ is a library that implements *ciphart* very closely to this paper, without much fluff. this should make integrating *ciphart* into other systems more convenient.

ciphart⁵ is an application for encrypting and decrypting files that makes use of **libciphart**. this application is intended for use by end-users or scripts, henceforth it has some fluff to treat mankind with dignity.

¹mail: toraboracaveman [at] protonmail [dot] com.

²<https://github.com/P-H-C/phc-winner-argon2>

³<http://www.tarsnap.com/scrypt/scrypt.pdf>

⁴<https://github.com/Al-Caveman/libciphart>

⁵<https://github.com/Al-Caveman/ciphart>

paper's layout

1	background	1
2	caveman's entropy	2
2.1	recursive hash	2
2.2	memory-hard hash	2
3	"perfect lie" theorem	2
4	"cacheable keys" discovery	3
4.1	why does it work	3
4.2	potential adversary strategies	4
5	application scenarios	4
5.1	a currently-useful scenario	4
5.2	a later-useful scenario	4
6	ciphart	5
6.1	parameters	5
6.2	internal variables	5
6.3	output	5
6.4	steps	5
7	parallelism	5
8	memory-hardness	5
9	memory-hardness	5
10	comparison	6
A	donations	6
A.1	bitcoin	6

1 background

we've got password p with $H(p)$ many shannon's entropy bits worth of information in it. so what does this mean?

fundamentally, it means that, on average, we'd need to ask $H(p)$ many perfect binary questions⁶ in order to fully resolve all ambiguities about p ; i.e. to fully get every bit of p .

but people use it to do less orthodox things, such as quantifying the amount of security p has against, say, brute-forcing attacks.

say that we've got a $8V$ bit key $k \leftarrow \text{hash}(p||s, 8V)$, derived from password p , where s is a salt. say that the attacker has s and k but wants to figure out p . in this case, he will need to brute-force the password space in order to find p that gives k . his cost is:

$$2^{H(p)} \left(\text{cost}(\text{hash}) + \text{cost}(\text{if } \hat{k} = k) \right) \quad (1)$$

definition 1. *the security of a system is the cost of the cheapest method that can break it.*

⁶one which, if answered, and on average, gets the search space reduced in half.

one way to estimate **cost** is to survey the asics industry. by surveying the asics industry to get an idea how much money it costs to get a given key, or password, space brute-forced within a target time frame⁷. this has an expensive housekeeping and is usually not possible to get any guarantees as we don't know about state-of-art manufacturing secrets that adversaries may have.

another way is to ignore anything that has no cryptographic guarantee. so, in (1), cryptography guarantees⁸ that $2^{H(p)}$ many **hash** calls are performed and that many equality tests. the **hash** call needs to be done once, so let's give it a unit of time 1. the equality test also needs to be called once, but since since it's so cheap it's easier to just assume that its cost is free. this way (1) becomes just:

$$2^{H(p)}(1 + 0) = 2^{H(p)} \quad (2)$$

further, for convenience, it seems that people report it in the \log_2 scale. i.e. $\log_2 2^{H(p)} = H(p)$. i think this is why people use password entropy as a measure of its security. not because it is the quantity of security, but rather because its the quantity of *simplified* security.

i like this shannon's entropy-based simplified security quantity, so i'm going to build on it.

2 caveman's entropy

2.1 recursive hash

if the **hash** function is replaced by an N -deep recursion over **hash**, like:

$$\begin{aligned} & \text{rhash}(p\|s, 8V, N) \\ &= \text{hash}(\text{hash}(\dots \text{hash}(p\|s, 8V), \dots, 8V), 8V) \end{aligned}$$

then, if **hash** is not broken, (1) becomes:

$$2^{H(p)} \left(N \text{cost}(\text{hash}) + \text{cost}(\text{if } \hat{k} = k) \right) \quad (3)$$

(2) becomes:

$$\begin{aligned} 2^{H(p)}(N + 0) &= N 2^{H(p)} \\ &= 2^{H(p) + \log_2 N} \end{aligned} \quad (4)$$

and the \log_2 -scaled becomes $H(p) + \log_2 N$.

at this point, thanks to cryptographic guarantees concerning properties of hashing functions, there is absolutely no security distinction from adversary's point of view between a password with shannon's $H(p) + \log_2 N$ entropy bits, and a password with just $H(p)$ entropy bits that made use of the N -deep recursive calls of **hash**.

shannon's entropy of p remains $H(p)$, but thanks to the recursive calls of **hash**, that password will be as expensive as another password \hat{p} , such that $H(\hat{p}) = H(p) + \log_2 N$.

i think it will be simpler if we introduce the function-dependent caveman's entropy C as a measure. it goes like this:

$$C(p, \text{hash}(\dots)) = H(p) \quad (5)$$

$$C(\hat{p}, \text{hash}(\dots)) = H(p) + \log_2 N \quad (6)$$

$$\begin{aligned} C(p, \text{rhash}(\dots, N)) &= H(p) + \log_2 N \\ &= H(\hat{p}) \end{aligned} \quad (7)$$

security-wise, there is no distinction between the more complex password \hat{p} , and the simpler password p that used **rhash**(..., N). so i really think we need to measure password security in C instead of H .

2.2 memory-hard hash

let **mhash** be like **rhash**, except that it also requires M many memory bytes such that, as available memory is linearly reduced from M , penalty in cpu time grows exponentially. let M be requested memory, A be available memory, and $e(M - A)$ be the exponential penalty value for reduction in memory, where $e(0) = 1$.

$$\begin{aligned} & \text{cost}(\text{mhash}(p\|s, N, M)) \\ &= \text{cost}(\text{rhash}(p\|s, N))^{e(M-A)} \end{aligned} \quad (8)$$

if **hash** in (1) is replaced by the M -bytes memory-hardened N -deep recursion hash function **mhash**, then (1) becomes:

$$2^{H(p)} \left(N^{e(M-A)} \text{cost}(\text{hash}) + \text{cost}(\text{if } \hat{k} = k) \right) \quad (9)$$

(2) becomes:

$$\begin{aligned} 2^{H(p)}(N^{e(M-A)} + 0) &= N^{e(M-A)} 2^{H(p)} \\ &= 2^{H(p) + \log_2 N^{e(M-A)}} \\ &= 2^{H(p) + e(M-A) \log_2 N} \end{aligned} \quad (10)$$

and caveman's entropy becomes:

$$C(p, \text{mhash}(\dots, N, M)) = H(p) + e(M - A) \log_2 N \quad (11)$$

3 “perfect lie” theorem

let p be a password with $H(p)$ shannon's entropy bits. let \hat{p} be a more complex password with $H(p) + e(M - A) \log_2 N$ shannon's entropy bits, where M , A and N are all positive numbers.

then caveman's entropy says that the following keys are information theoretically indistinguishable for as long as only p and \hat{p} remain unknown (everything else is known, such as the distribution from which p and \hat{p} was sampled), and for as long as **hash** is not broken:

⁷see the *scrypt* paper for an example.

⁸statistically by confidence earned through peer review and attempts to break encryption algorithms.

- $k \leftarrow \text{mhash}(p||s, N, M)$
- $\hat{k} \leftarrow \text{hash}(\hat{p}||s)$

in other words:

$$C(p, \text{mhash}(\dots, N, M)) = H(\hat{p}) \quad (12)$$

since the assumption that passwords are kept away from the adversary is fundamental in a symmetric encryption context, i think it makes sense that we measure our security with memory-hard key derivation functions using the caveman's entropy C instead of shannon's entropy H .

from a security point of view, it will feel absolutely identical to as if the password got injected with extra shannon's entropy bits. no one can tell the difference for as long as the fundamental assumption of hiding passwords is honoured, as well as the hashing function `hash` is not broken.

in other words, we can say, if password p is unknown, and `hash` is not broken, then we have injected into p extra shannon's entropy bits. this lie will be only discovered after p is revealed.

if you think that it is impossible for this *lie* to be *truth* under the secrecy of p , then i've done an even better job: proving that cryptographically secure hashing functions do not exist. likewise, same can be trivially extended to: cryptographically symmetric ciphers do not exist.

so you have to pick only one of these options:

1. either accept that the lie is truth. i.e. accept that we've injected shannon's entropy bits into p , for as long as only p is not revealed.
2. or, accept that cryptographically-secure hashing and symmetric-encryption functions do not exist.

theorem 1 (perfect lie⁹). *for as long as password p remains a secret and `hash` is not broken, derived keys $k_1 = \text{rhash}(p, \dots, N)$ and $k_2 = \text{mhash}(p, \dots, N, M)$ will have their shannon's entropies, $H(k_1)$ and $H(k_2)$, increase beyond $H(p)$ in order to equate their caveman's entropies $C(p, \text{rhash}(\dots, N))$ and $C(p, \text{mhash}(\dots, N, M))$, respectively.*

the reason this lie is appealing is because it simplifies our quantification of the amount of security that we have gained by using a given key derivation function, such as `rhash` or `mhash`.

without treating this lie as truth, our only hope would be surveying the asics industry. but with this lie, we have one more approach to get a feel of the gained security quantity by just accepting caveman's entropy C as shannon's entropy H , and move on as if the lie is truth, and no one can notice it.

we can also look at it from the perspective of *occam's razor*. i.e. if two things are not distinguishable from one

another, then assuming that they are just the same thing is simpler than assuming otherwise.

to be more specific about *occam's razor*: (1) each assumption bit has a positive probability of error by definition, (2) since assuming that indistinguishable things are different than one another is more complex (i.e. more assumption bits) than assuming not, and (3) since there is no observable difference between the two things, therefore it necessarily follows that our model's total error will be reduced if we accept that the indistinguishable things are identical (i.e. which is what theorem 1 says).

4 “cacheable keys” discovery

discovery 1 (cacheable keys). *caching keys securely is easily doable, and great security utility exists in doing so for expensively-derived keys.*

4.1 why does it work

- when expensively derived keys are cached, only the first key derivation call will be expensive, while subsequent calls will be semi-instantaneous. this effectively allows users to tolerate much more expensive, or secure, key derivation as it only happens during the initial, say, login phase. subsequent use of the extremely expensive key is instantaneous.

so instead of having the user use a somewhat expensive key derivation by waiting say, 3 seconds in each login, he will —instead— wait, say 10 seconds in his initial login in order to utilise a much more expensive key derivation, and then wait near 0 seconds for every subsequent login as the expensive key is cached.

- derived keys can be cached securely, without increasing most users' assumptions. e.g. cached keys can live in a *dm-crypt* partition that is encrypted with a large encryption key that is stored properly, and the cache can have strict read permissions so that only the unique user that runs `ciphart` executable can read it.

this only requires to trust the user `root`, which is already trusted by almost everyone. so we are not introducing a new assumption.

for most people, if `root` is compromised, then the adversary can break every other key derivation function, including those that do not cache keys, by simply, say, running a keylogger.

so, practically, we are not increasing the assumptions, but we are only increasing the value that we can extract from the assumptions that we already have.

most importantly, utilising discovery 1 allows us to achieve a memory-harder key derivation in an extremely usable way. more details on memory-hardness later.

⁹i call theorem 1 *perfect lie* in a sense that a perfect lie is indistinguishable from truth.

4.2 potential adversary strategies

let's see what may the adversary try to do against a key caching system:

- **adversary strategy 1:** hack into user's system and execute a program as that user that tries to read the password cache file to obtain the memory-harder key inside it.

answer: he will not be able to read the file due to strict read permissions of the cached files as set earlier.

- **adversary strategy 2:** steal user's hard disk and try to mount it in his system, to login as root and change permissions of the key cache files.

answer: the partition where the cached files are saved are properly encrypted, so he can't see the cached files, let alone changing their read permissions.

- **adversary strategy 3:** break into machine's root account.

answer: he will succeed, but then he can also run a keylogger, which will also break every other key derivation function, even those who do not use key caching, such as *script*, *argon2*, etc.

5 application scenarios

5.1 a currently-useful scenario

user has a password manager which generates unique 256 bit entropy keys for each online service that he uses. the user also renews keys for his online services every now and then. so his online accounts generally have high security.

but user's problem is how to lock and unlock his password manager's passwords database. should he use a physical usb-stick key that types a high-entropy key that encrypts and decrypts the database? the user doesn't want this physical key because of several reasons:

- he tends to lose his keys a lot, and, for certain tasks, the risk of needing to wait for until he gets a backup usb-stick key is too much.
- he doesn't want to be caught having cryptographic usb-stick keys, because as such is an evidence that he has encrypted content. the user wants to have the choice to lie that he has no clue. so not having usb-stick keys with him helps his case to lie.
- he wants to be torture-resistant so that an adversary cannot forcefully take his keys from him in order to login into his services. he may rather want to die than to give the password to the adversary. this works because, so far, the brain is a pretty private information store.

in this scenario, the user memorises a sensible password that he can remember, with enough initial entropy, and

then uses *ciphart*, preferably with disk caching, to inject large amounts of entropy bits into his derived keys, way beyond the reach of pre-*ciphart* key derivation functions; thanks to theorem 1 and discovery 1.

here, the expensively derived key is only used to unlock a local password manager, which offers a protection against situations where a backup copy of the passwords database is stolen. this may enable the user to store his passwords manager in an online file synchronisation service for more convenient system migrations to further reduce his login delays should he face the need to migrate to a new, say, personal computer.

advertisement 1. *in case you're interested in such a passwords manager, i've also made nsapass¹⁰ — a flexible and a simple passwords manager in a few hundreds of python lines of code, that uses ciphart by default. i think this is the best command-line interfacing passwords manager by far, for its usability, and for the fact that auditing it is easy, thanks for it being only in a few hundreds of python lines.*

5.2 a later-useful scenario

all input password fields will internally call `libciphart` to derive more expensive keys. this way, applications, such as *firefox*, *mutt*, ..., will never send actual passwords, but will only send *ciphart*-derived keys with increased shannon's entropy content.

thanks to theorem 1, this will automatically increase shannon's entropy of all users password without requiring users to memorise harder passwords. thanks to discovery 1, the user will not face any delay in his daily use.

i also think that it is *generally* better to have expensive key derivation functions in the client side as opposed to the server side. because remote servers always have the incentive to reduce the complexity of the key derivation function in order to free more resources for other things that bring them money.

either way, *ciphart* is perfectly usable on the server side. it is just that i think *ciphart*, *argon2*, *script*, ... make better sense when placed on the client side.

¹⁰<https://github.com/Al-Caveman/nsapass>

6 ciphart

6.1 parameters

P	password.
S	salt.
M	total random-access memory in bytes.
D	total hard-disk memory in bytes.
F	temporary file's path.
Y	whether key caching is enabled.
L	number of memory lanes for concurrency.
T	number of tasks per lane segment.
H	number of lane segments per hard-disk read.
B	minimum <i>caveman's entropy bits</i> to inject into p .
K	output key's size in bytes.

6.2 internal variables

enc	encryption function.
hash	hashing function.
read	hard-disk file reading function, with seeking. e.g. read (x, y, z) reads z many bytes from file x after seeking y bytes forward.
write	hard-disk file writing function.
C	$\leftarrow \begin{cases} 64 \text{ bytes} & \text{if } \mathbf{enc} \text{ is } \mathit{xchacha20} \\ 16 \text{ bytes} & \text{if } \mathbf{enc} \text{ is } \mathit{aes} \\ \dots & \end{cases}$ this to reflect the block size of the encryption algorithm that implements enc .
V	$\leftarrow \begin{cases} 32 \text{ bytes} & \text{if } \mathbf{enc} \text{ is } \mathit{xchacha20} \\ 16 \text{ bytes} & \text{if } \mathbf{enc} \text{ is } \mathit{aes-128} \\ 32 \text{ bytes} & \text{if } \mathbf{enc} \text{ is } \mathit{aes-256} \\ \dots & \end{cases}$ this is the size of the encryption key that's used to solve <i>ciphart</i> 's tasks. this is different than the enc -independent K which is possibly used by other encryption algorithms in later stages ¹¹ .
\hat{T}	$\leftarrow \max(\lceil VC^{-1} \rceil, T)$. this is to ensure that we have enough encrypted bytes for new keys.
\hat{T}	$\leftarrow \hat{T} - (\hat{T} \bmod 2) + 2$. this is to ensure that there is an even number of tasks in a segment. why? because we need a buffer for storing the clear-text and another for storing the output cipher-text.
\hat{M}	$\leftarrow M - (M \bmod C\hat{T}L) + C\hat{T}L$. this is to ensure that it is in multiples of $C\hat{T}L$. why? so that all segments are of equal lengths in order to simplify <i>ciphart</i> 's logic. e.g. it wouldn't be nice if the last segments were of unequal sizes.
G	$\leftarrow \hat{M}C^{-1}\hat{T}^{-1}L^{-1}$. total number of segments per lane.

¹¹at the expense of losing the meaning of *caveman's entropy bits*.

n	$\leftarrow 0$. actual number of times enc is called.
m_i	C -bytes memory for i^{th} task in the \hat{M} -bytes pad.
n_l	$\leftarrow lG\hat{T}$. nonce variable for l^{th} lane with at least 64 bits.
f	$\leftarrow 0$. a counter indicating number of times memory is filled with \hat{M} many bytes.
d	$\leftarrow 0$. a counter indicating total number of saved blocks into hard-disk.
h	$\leftarrow 0$. a counter indicating number of processed lane segments since the last hard-disk read.
q	$\leftarrow 0$. a counter indicating number of times key was updated from the hard-disk.
$*v$	$\leftarrow *hash(P\ S\ M\ D\ \dots\ K, V)$. a pointer to the first byte where V -bytes key is stored. v is the key itself.
Z	$\leftarrow hash(v\ 0, V)$. file name where output key, k , is expected to be cached, if k was previously cached.

6.3 output

k	K -bytes key.
n	total number of times enc was actually called. $\log_2 n$ is total number shannon's entropy bits that <i>ciphart</i> injected into k , such that $\log_2 n \geq B$.

6.4 steps

steps of *ciphart* is shown in algorithm 1. this corresponds to *argon2d*. adding a *ciphart-i* variant is a trivial matter, i just didn't do it yet because my threat model currently doesn't benefit from a password independent variant.

7 parallelism

since iterations of the loop in line 6 in algorithm 1 are fully independent of one other, they can quite happily utilise L cpu cores, specially when segment sizes, T , are larger.

8 memory-hardness

Proof. algorithm 1 is just a variation of *argon2d*, except that it uses an encryption function, **enc**, instead of a hash-function. so if *argon2d* is memory-hard, then so is *ciphart*. \square

9 memory-hardness

thanks to discovery 1, memory-hardness is possible. this process goes like this:

1. starts by recursively encrypting D many bytes of some predefined sequence, such as zeros or `0xdbdb...`, using a key derived from an cheaply-hashed password p . by the end of this step, we have D many key-based random sequence saved in the an encrypted hard disk partition.

algorithm 1: ciphart

```
1 if  $Y$  and exists( $Z$ ) then
2    $k, n \leftarrow \text{read}(Z, 0, V + \text{sizeof}(n));$ 
3   return  $k, n$ 
4 while 1 do
5   for  $g \leftarrow 0, 1, \dots, G - 1$  do
6     for  $l \leftarrow 0, 1, \dots, L - 1$  do
7       for  $t \leftarrow 0, 1, \dots, T - 1$  do
8          $i \leftarrow gLT + lT + t;$ 
9         if  $t < T - 1$  then
10           $j \leftarrow i + 1;$ 
11        else if  $t = T - 1$  then
12           $j \leftarrow i - T + 1;$ 
13         $m_j \leftarrow \text{enc}(m_i, n_l, *v);$ 
14         $n_l \leftarrow n_l + 1;$ 
15        if  $f = 0$  then
16           $*v \leftarrow m_j \bmod (gLTC + tC - V);$ 
17          if  $*v \geq gLTC - V$  then
18             $*v \leftarrow *v + lTC;$ 
19        else
20           $*v \leftarrow m_j \bmod (\hat{M} - LTC + tC - V);$ 
21          if  $*v \geq gLTC + tC - V$  then
22             $*v \leftarrow *v + LTC;$ 
23          else if  $*v \geq gLTC - V$  then
24             $*v \leftarrow *v + lTC;$ 
25       $n \leftarrow n + LT;$ 
26      if  $d \leq D$  then
27        for  $i \leftarrow gLT, \dots, gLT + (T - 1)$  do
28          write( $F, m_i$ );
29           $d \leftarrow d + 1;$ 
30          if  $d \geq D$  then
31            break;
32      else
33         $h \leftarrow h + L;$ 
34        if  $h \geq H$  then
35           $*v \leftarrow \text{read}(F, v \bmod (d - V), V);$ 
36           $q \leftarrow q + 1;$ 
37           $h \leftarrow 0;$ 
38        if  $f \geq 2$  and  $q \geq 1$  and  $n \geq 2^B$  then
39           $g_{\text{last}} \leftarrow g;$ 
40          go to line 42;
41     $f \leftarrow f + 1;$ 
42   $i \leftarrow g_{\text{last}}LT;$ 
43   $k \leftarrow \text{hash}(m_{i+0T} \| m_{i+1T} \| \dots \| m_{i+(L-1)T}, K);$ 
44  if  $Y$  then
45    write( $Z, v \| n$ );
46  delete( $F$ );
47  return  $k, n$ 
```

2. update the derived key to be the last V bytes in the D bytes file.
3. move on to run a variant of *argon2*, except that, every time S many random-access memory lane segments are completed, some V many bytes from D are chosen randomly, not only based on the content the random-access memory, but also based on the D bytes. this step makes *ciphart* require $D + M$ bytes.
4. delete the D bytes, just to free up the disk space.

if S is large enough, this can be done efficiently without blocking the cpu noticeably, as the randomly obtained V bytes can be read by using the $O(1)$ operation `seek` over the D bytes.

10 comparison

A donations

this work is sponsored by a bright nerdiness that has unexpectedly sparkled somewhere in an endless dark space. nothing here was supposed to happen. the cause remains unknown, so we call it *something random with large entropy*.

ancient legend has it that, every time a donation is made for a good cause, a beautiful torch is kindled somewhere deep within an otherwise cold and dark space.

the torch is pure with an innocent smile every time a spending is committed towards good. but, she drops a tear when the spending is done in excess. despite her attempts to maintain a steady light, sometimes her tears force her light to flicker, and sometimes the flickering causes her to disappear. when she is gone, she is never back again, except in the memories of those who have once witnessed her charm...

so, if you opt for a donation, i'd appreciate exercising wisdom, so that the sensible balance is not exceeded; lest we want a beautiful torch to be born, only to make her disappear in her tears.

A.1 bitcoin



bc1qtyljztgd0yu4v7f8hyfzufn7nu692v9fc88jln