

Analysis of the WinZip Encryption Method

Paper by: Tadayoshi Kohno

Presented by: Ken, Mike, Jeremy & Paul



New WinZip 9.0 Now with AES Encryption

The popular compression utility for Microsoft Windows computers

- “Easy-to-use AES encryption” - Advanced Encryption version two (AE-2)
 - Derives AES and HMAC-SHA1 keys from user’s passphrase
 - Encrypts compression output with with AES-CTR
 - Authenticates resulting ciphertext with HMAC-SHA1

A Secure Implementation?

- Proven secure MAC:
 - HMAC-SHA1
- Proven secure Encryption:
 - AES in counter mode
- Proven secure combination:
 - Encrypt-then-MAC





But...

“ Security products must be evaluated as a whole, and the security of a whole product may not follow as a simple corollary of the security of the underlying components.”

Compression and Encryption

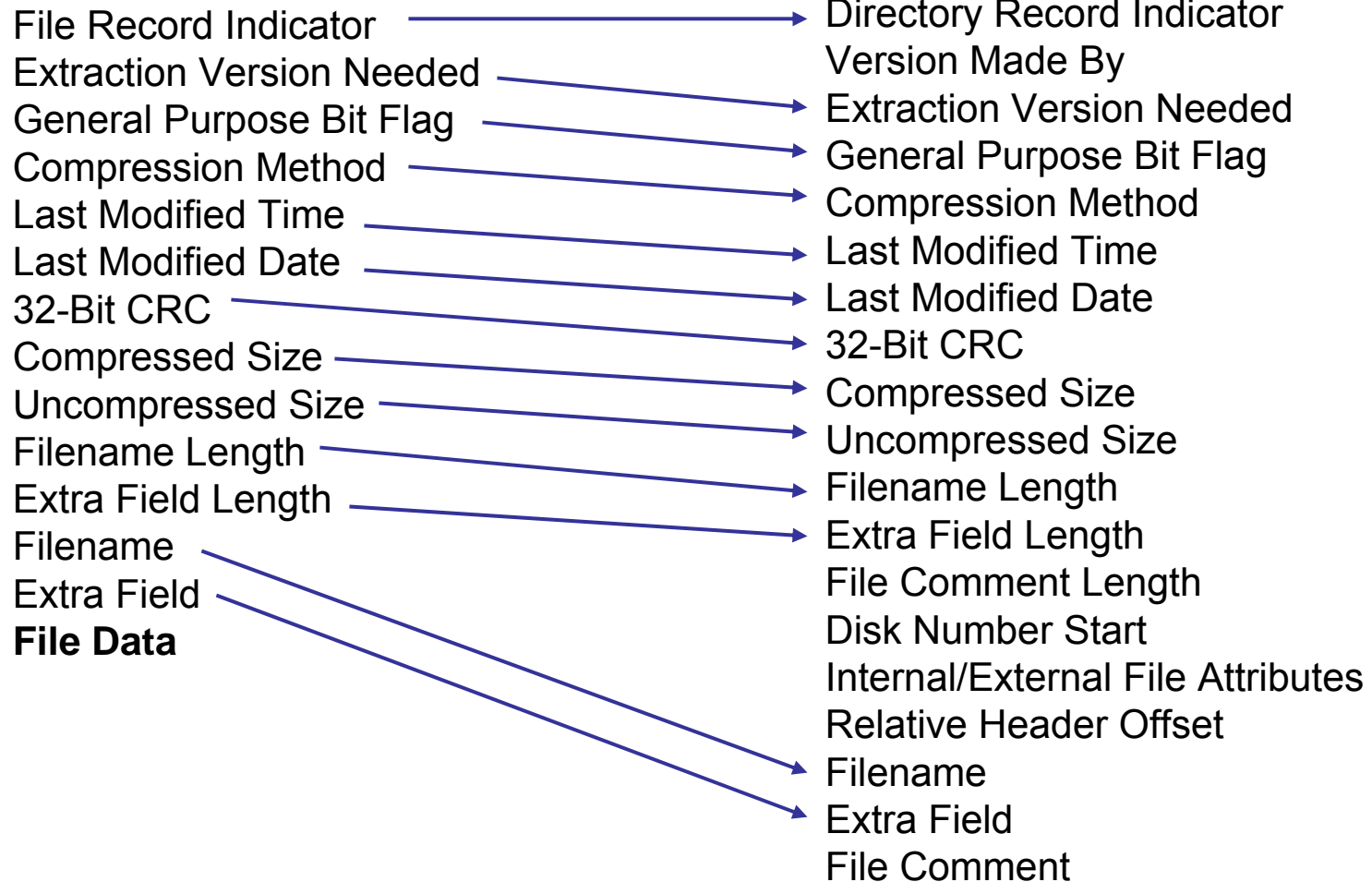
- WinZip creates two records for each file
 - Main file record
 - Central directory record
- Each Zip archive contains (in order):
 - The main file records concatenated together
 - The central directory records concatenated together
 - An End-of-Archive record

Note: A WinZip archive can contain multiple files. Each file is compressed/encrypted independently.

Archive Contents

The Main File Record

The Central Directory Record



Archive Contents

The Main File Record

File Record Indicator
Extraction Version Needed
General Purpose Bit Flag
Compression Method
Last Modified Time
Last Modified Date
32-Bit CRC
Compressed Size
Uncompressed Size
Filename Length
Extra Field Length
Filename
Extra Field
File Data

The Central Directory Record

Directory Record Indicator
Version Made By
Version Needed to Extract
General Purpose Bit Flag
Compression Method
Last Modified Time
Last Modified Date
32-Bit CRC
Compressed Size
Uncompressed Size
Filename Length
Extra Field Length
File Comment Length
Disk Number Start
Internal/External File Attributes
Relative Header Offset
Filename
Extra Field
File Comment

Important Archive Contents

The Main File Record

Compression Method
Last Modified Time
Last Modified Date
32-Bit CRC
Uncompressed Size
Filename
Extra Field
File Data

With AE-2 Encryption Enabled the File Data Field Contains:

Salt
Password Verification Value
Encrypted File Data
Authentication Code

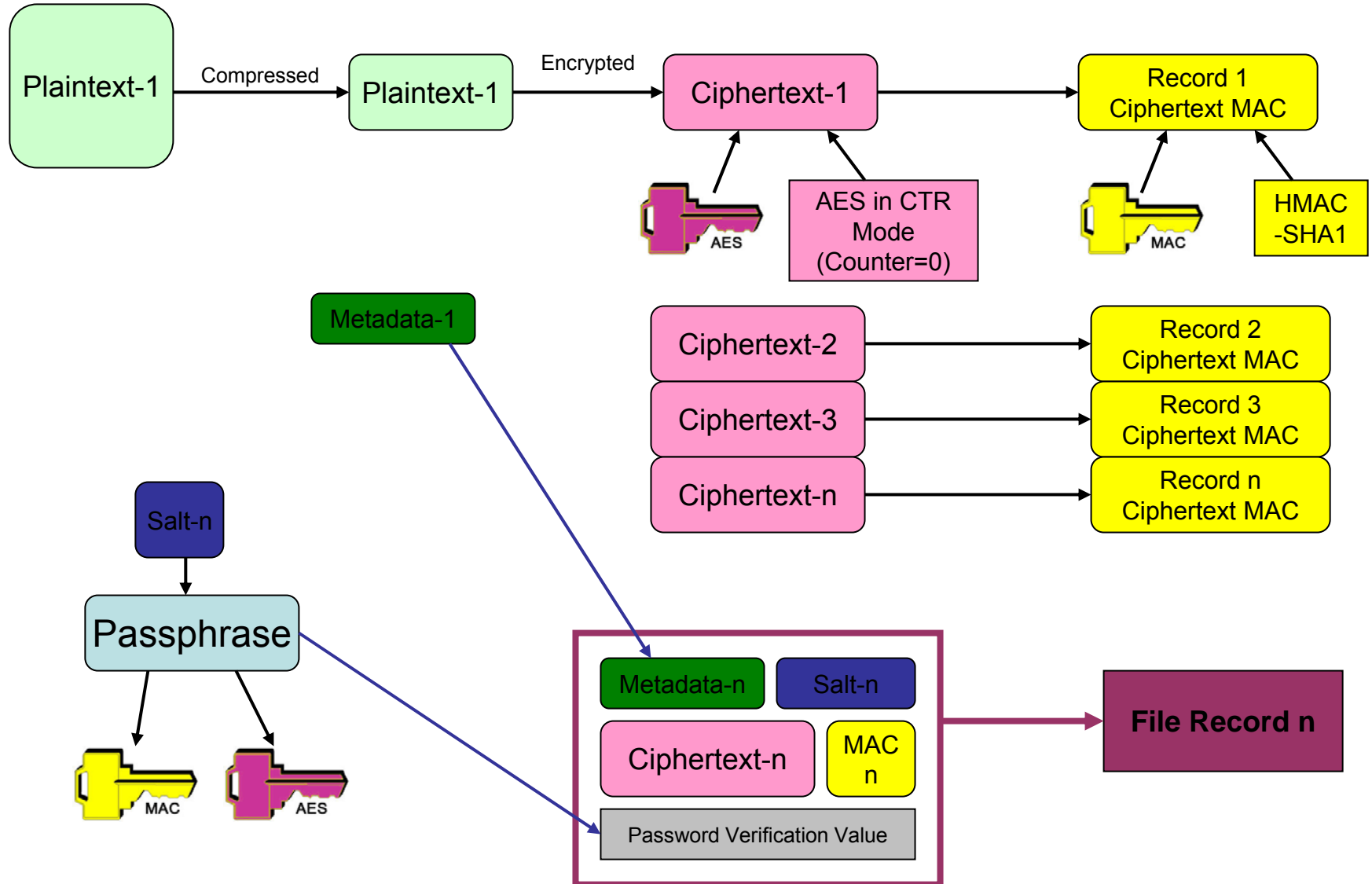
The Central Directory Record

Compression Method
Last Modified Time
Last Modified Date
32-Bit CRC
Uncompressed Size
Filename
Extra Field

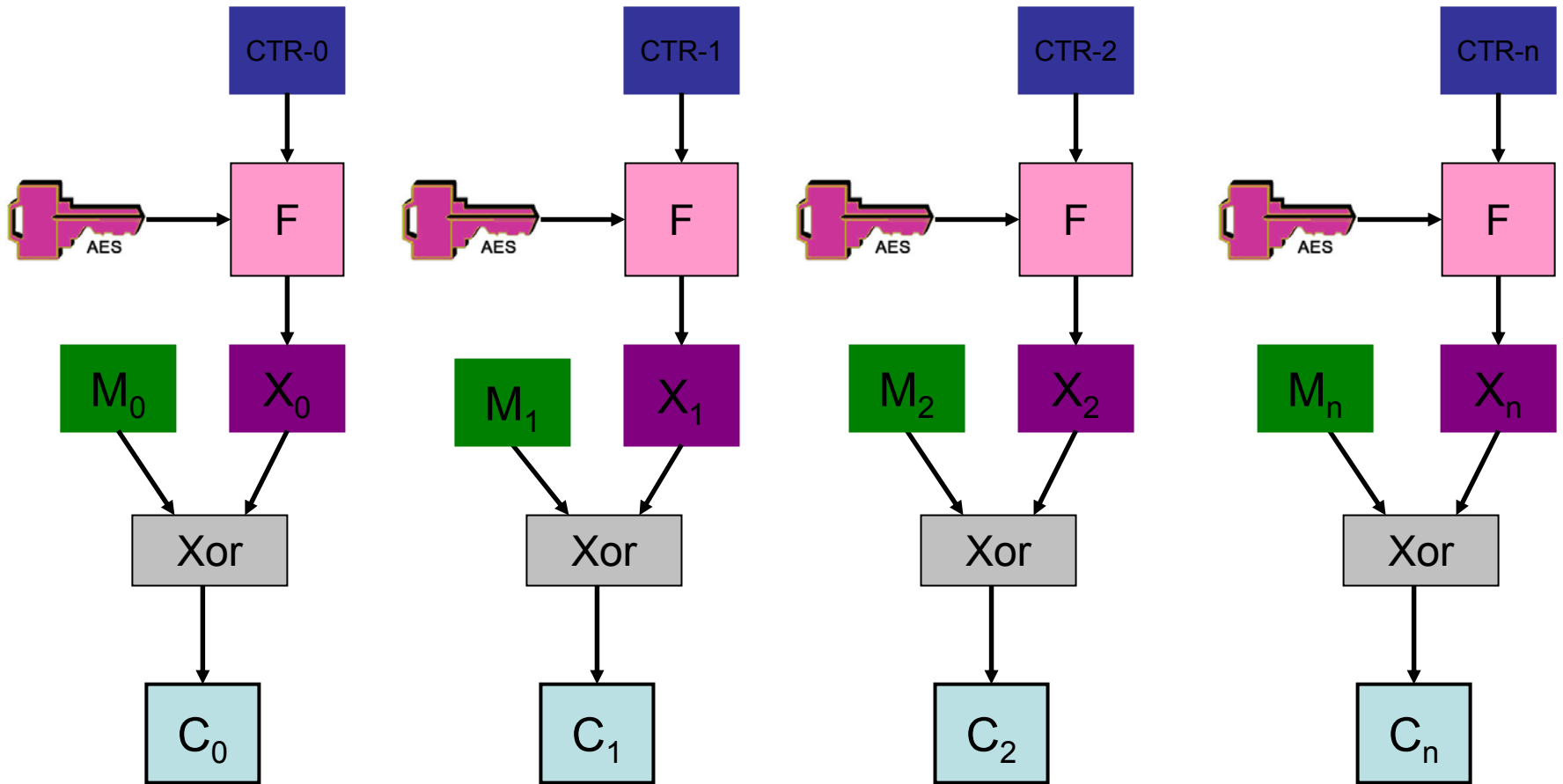
With AE-2 Encryption Enabled the Extra Fields Contain:

Extra Fields Header ID
Data Size
Version Number
Vendor ID
Encryption Strength
Actual Compression Method

File Encryption and Authentication Code Process



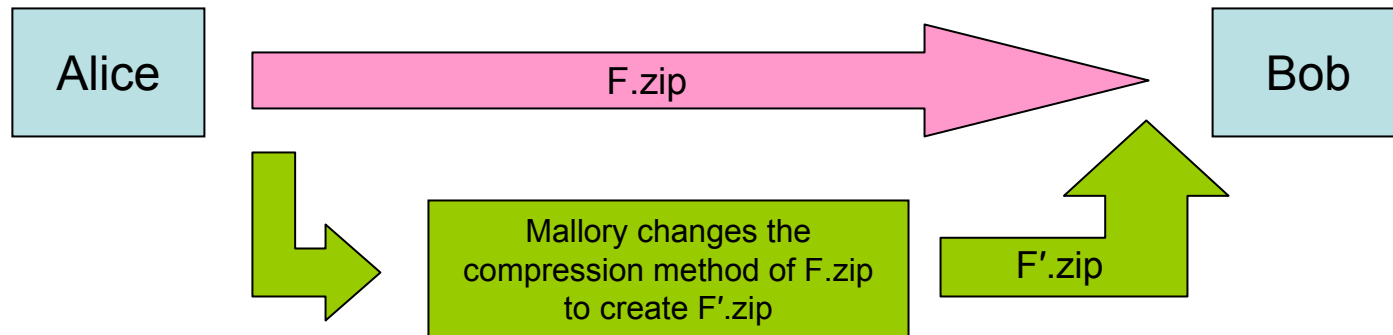
Counter Mode AES Encryption



WinZip Security Problems:

- ✓ Interactions between compression and the AE-2 encryption method.
- ✓ The names of files and their interpretations
- ✓ Information leakage from encrypted files' metadata
- ✓ Interactions with AE-1 and a chosen-protocol attack
- ✓ Archives with both encrypted and unencrypted files.
- ✓ Key collisions and repeated keystreams

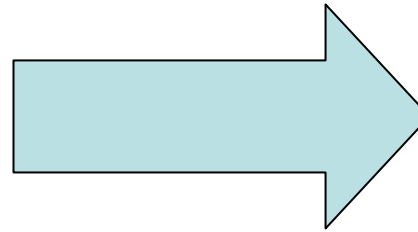
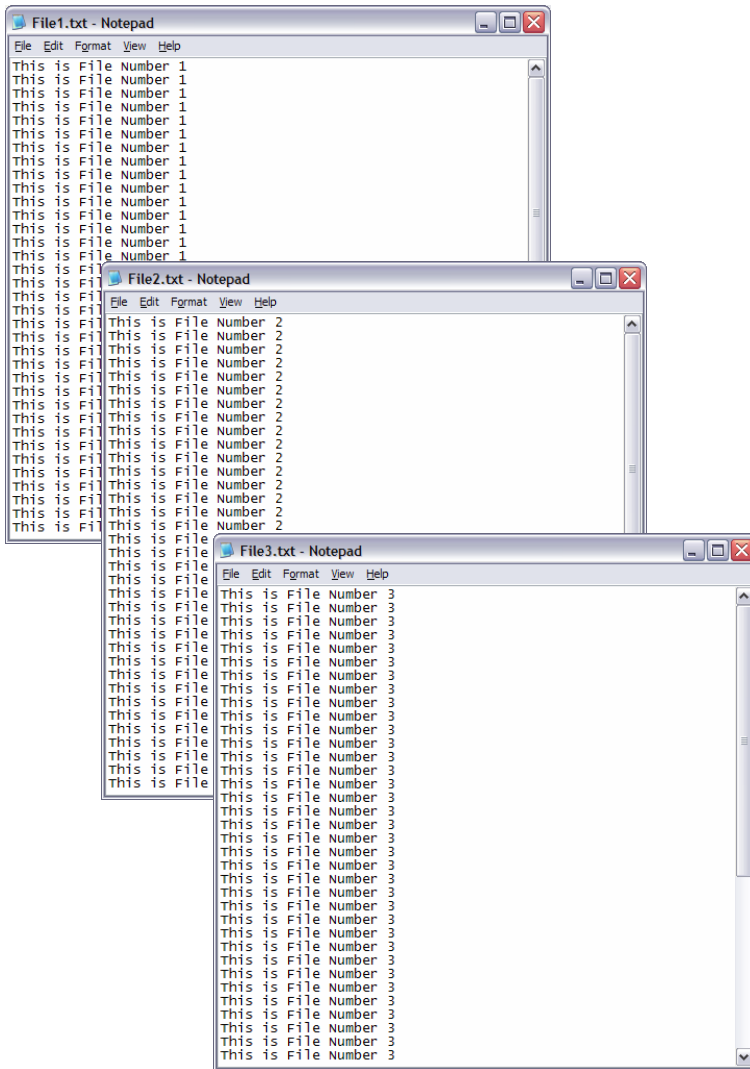
Exploiting the Interaction Between Compression and Encryption



Recall that the metadata is *not Authenticated*, therefore Mallory can change these values without voiding the HMAC-SHA1 tag.

- When Bob attempts to decrypt F'.zip (with the wrong compression method), the contents will be garbage.

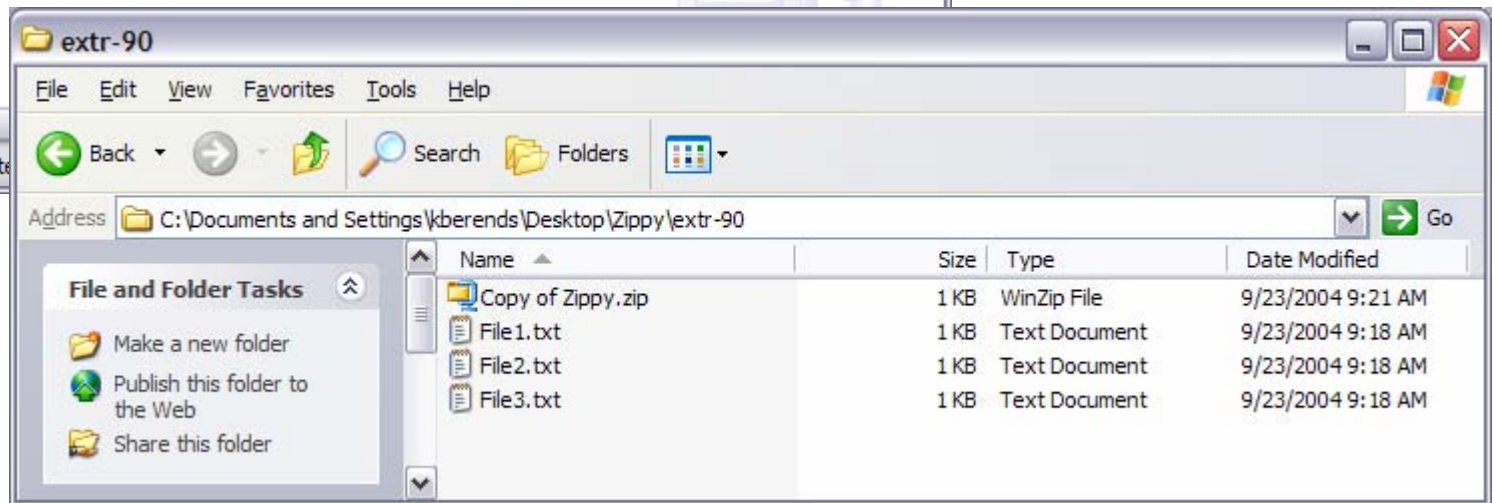
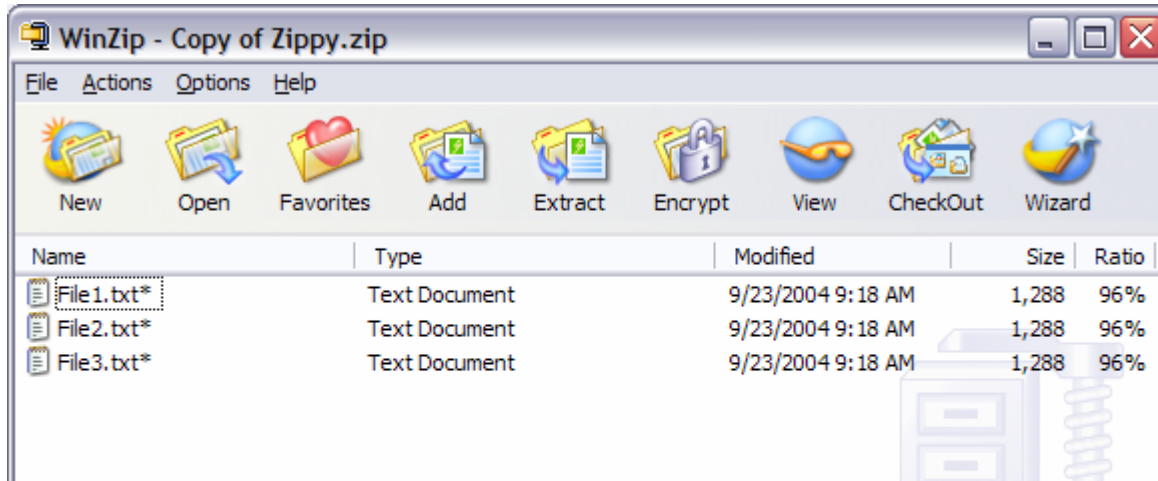
Create Encrypted Zip Archive



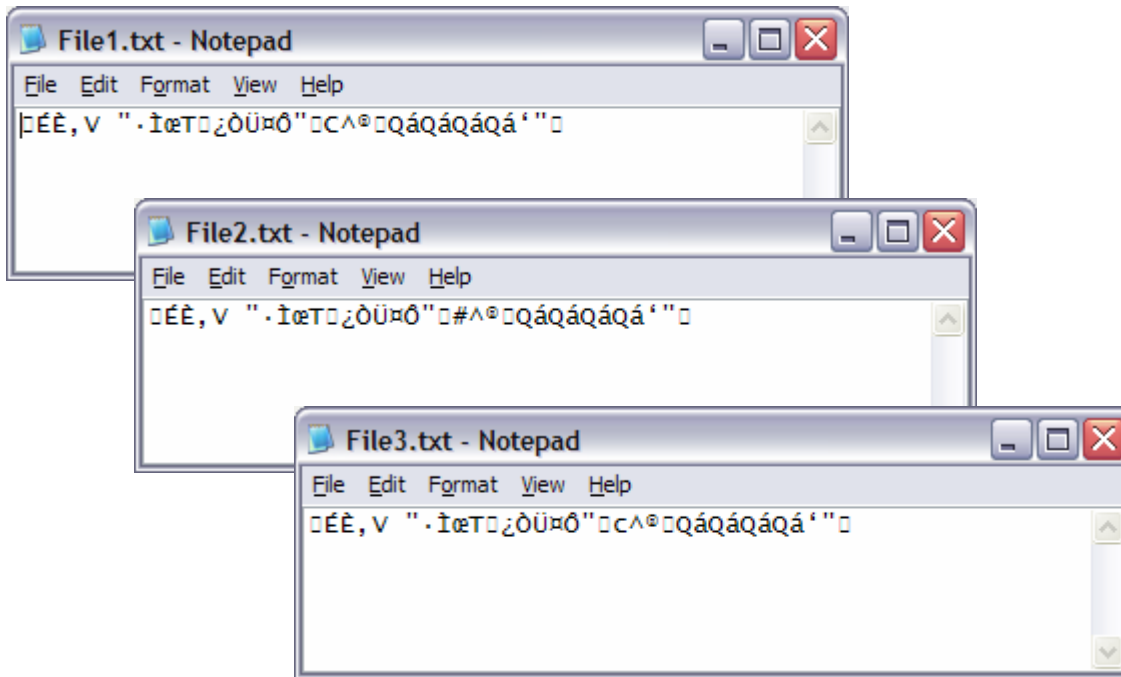
Zippy.zip

Using 128-bit AES
Encryption

Decryption of the Modified Archive - WinZip 9.0

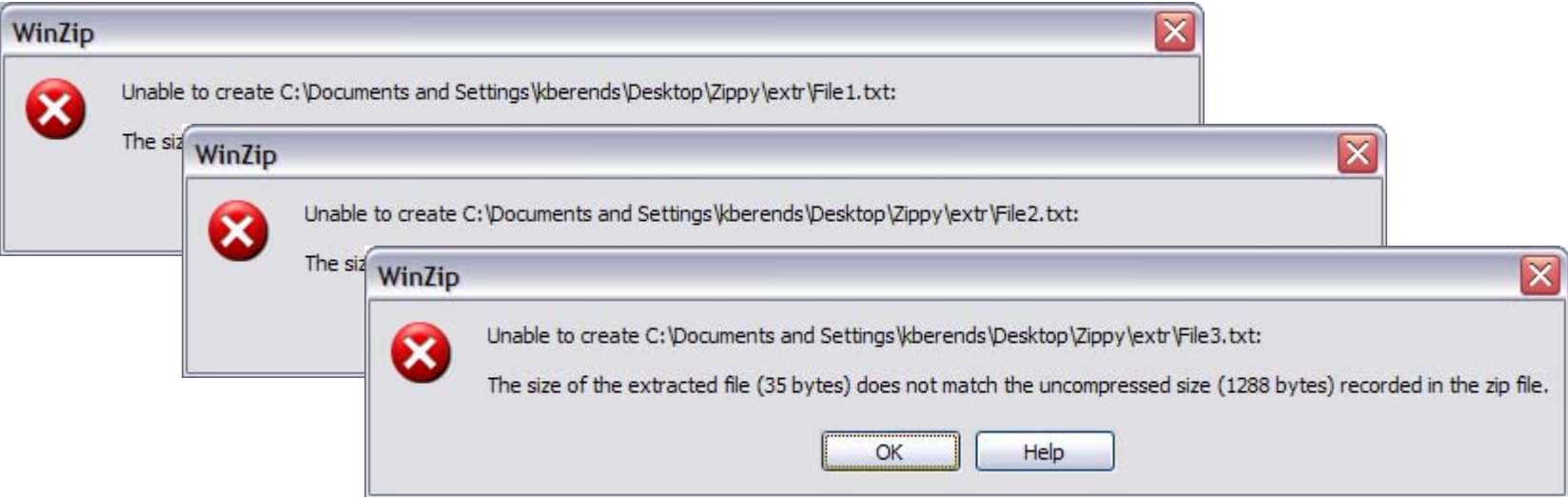
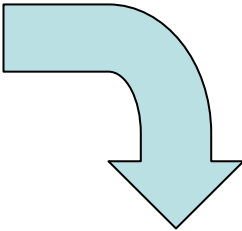
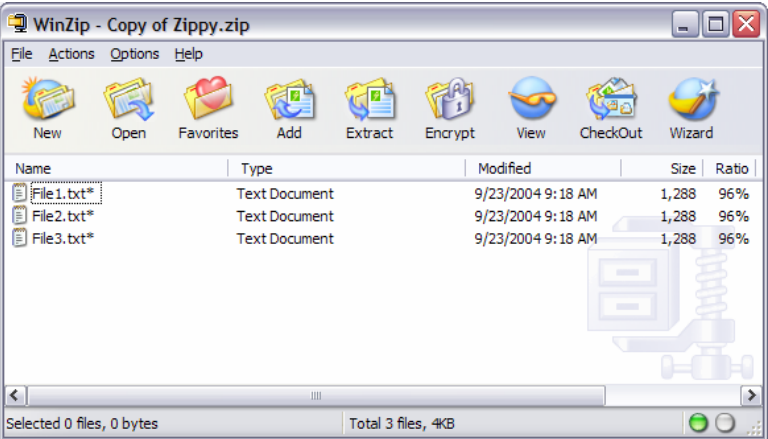


Garbage...



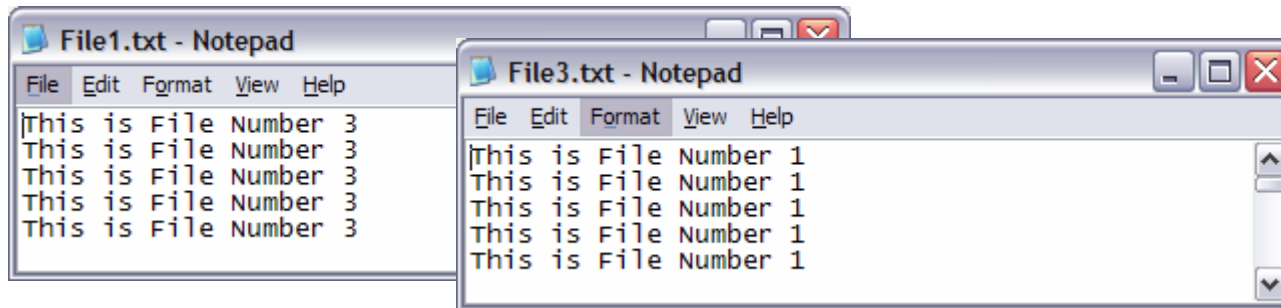
- If Mallory obtains this garbage, he can reconstruct F.zip.
- Is it practical for Mallory to obtain this garbage?

Decryption of the Modified Archive - WinZip 9.0 SR-1



Exploiting the Association of Applications to Filenames

- A variant of the previous scenario could also be to simply change the filename extension. (i.e. from .doc to .xls)
- Or the entire filename: Swap Alice-Salary.dat with Mallory-Salary.dat





Information Leakage

- *Cleartext Metadata:
 - Filenames, modified dates & times, CRC's, & file lengths
- Compression as a 'Side-Channel' (John Kelsey):
 - Compare original and compressed file sizes
 - Supplements pre-existing partial knowledge
 - Compare the compression ratios of related files
- Why?
 - Engineering or Design Complexities
 - Functionality
 - The ability to view archive contents without entering the passphrase

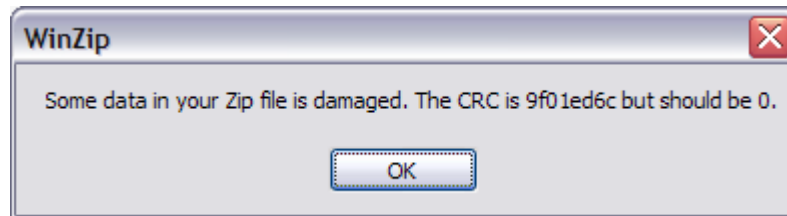
* The WinZip documentation notes the existence of such cleartext metadata, but does not address the security implications.

AE-1 vs. AE-2

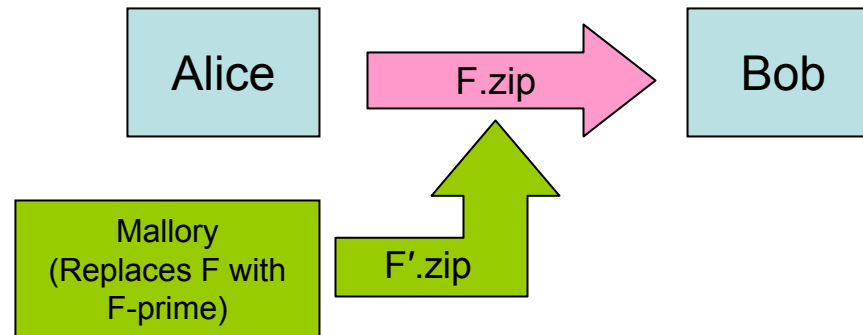
- Two methods of AES encryption used by WinZip.
- Due to a security flaw in AE-1 (CRC of plaintext is included in unencrypted format in the output), it was replaced by AE-2 in WinZip 9.0 Beta 3.
 - The CRC is a 32-bit checksum used to detect corrupted data.
- Backward compatibility is maintained, a little too well:
 - http://www.winzip.com/aes_info.htm
 - “Files encrypted using the AE-1 method *do* include the standard Zip CRC value. This, along with the fact that the vendor version stored in the AES extra data field is 0x0001 for AE-1 and 0x0002 for AE-2, **is the only difference** between the AE-1 and AE-2 formats.”
 - ZIP utilities that support AE-2 must support AE-1, and during decryption of AE-1 files, they should verify that the CRC matches.

Backwards compatibility exploited

- Adversary can force WinZip to use AE-1 decryption on an AE-2 encrypted file
 - Just change the vendor ID
 - Remember, everything else is the same: The same process is used to decrypt the file either way.
 - But now, WinZip will verify the CRC field!



Adversary can guess the content

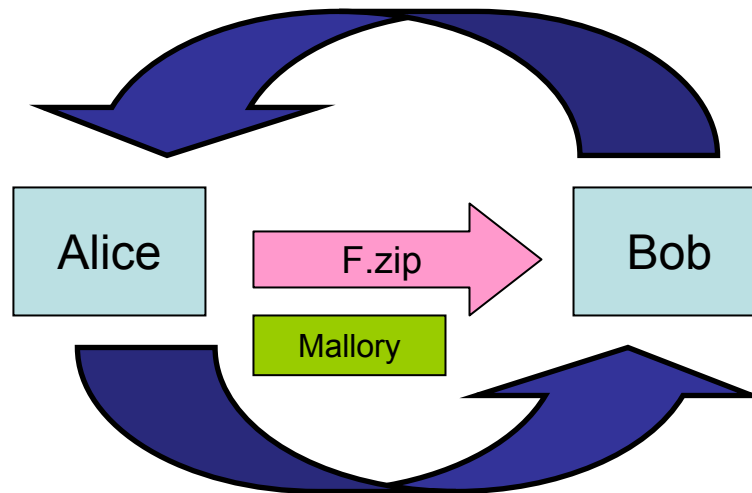


- Mallory guesses the content of a file in F.zip.
- Mallory computes the CRC of his guess.
- He then modifies F to F': changes ID number from AE-2 to AE-1 and inserts the CRC of his guess into the CRC field of the file.

Adversary can guess the content

- Bob receives F' .zip and attempts to decrypt it.
 - If Mallory's guess was correct, Bob will decrypt without errors.
 - If Mallory does not see a complaint, will assume guess was correct.
 - If Mallory's guess was incorrect, Bob will get an error, and complain to Alice.
 - Mallory will intercept this complaint.

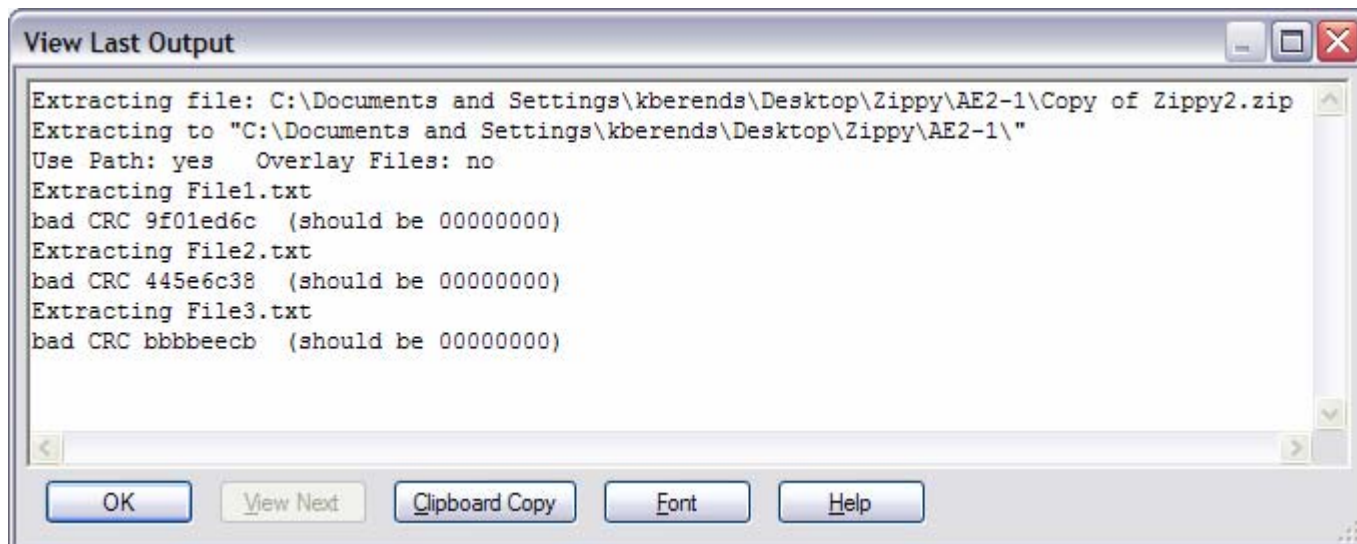
Adversary can guess the content



- Unfortunately, this is an online attack (requires active participation by Alice and Bob for each guess) – can we do better?

Offline attack

- If Bob includes WinZip's error messages or log files (with the CRC) in his complaint to Alice.
- Then, Mallory can intercept this complaint and conduct an offline attack.



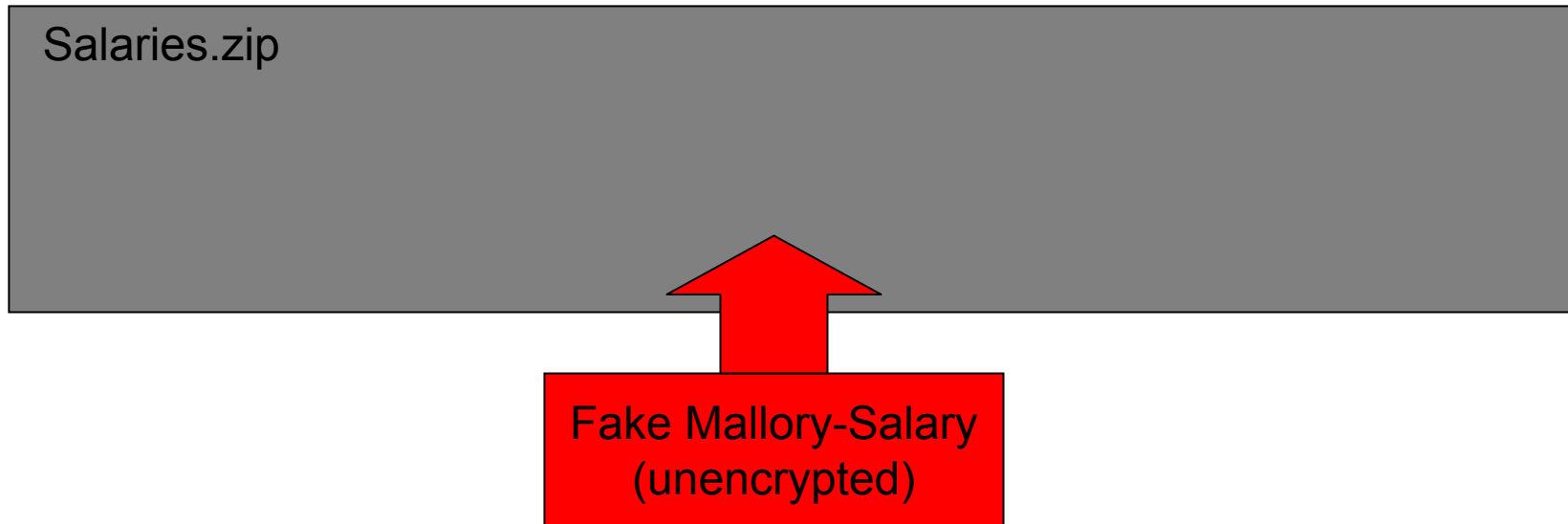
```
View Last Output
Extracting file: C:\Documents and Settings\kberends\Desktop\Zippy\AE2-1\Coppy of Zippy2.zip
Extracting to "C:\Documents and Settings\kberends\Desktop\Zippy\AE2-1\"
Use Path: yes  Overlay Files: no
Extracting File1.txt
bad CRC 9f01ed6c (should be 00000000)
Extracting File2.txt
bad CRC 445e6c38 (should be 00000000)
Extracting File3.txt
bad CRC bbbbeecb (should be 00000000)
OK View Next Clipboard Copy Font Help
```

Attacking Zip Encryption at the File Level



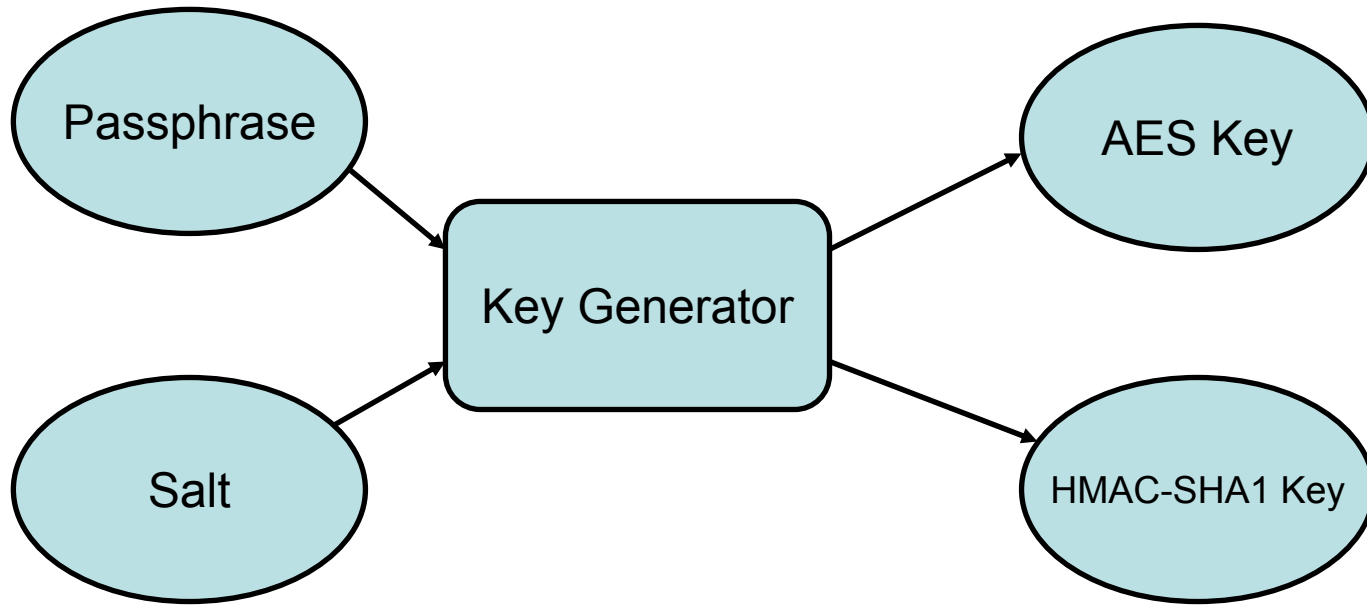
- Each file within the archive is encrypted separately
- Not all files within the archive may be encrypted – some might just be compressed
- Attacker can replace individual encrypted files with unencrypted files containing any content

Attacking Zip Encryption at the File Level ⁽²⁾



- Bob will receive no warning when decrypting.
- Usability issue?

Dictionary Attacks



- Salt (random value) used to impede dictionary attacks
 - Intended to prevent attacker from pre-computing associations between passphrases and keys.

How do we fix these issues?



Fixing Compression/Encryption Methods and File Names/Associations

Authenticate everything!

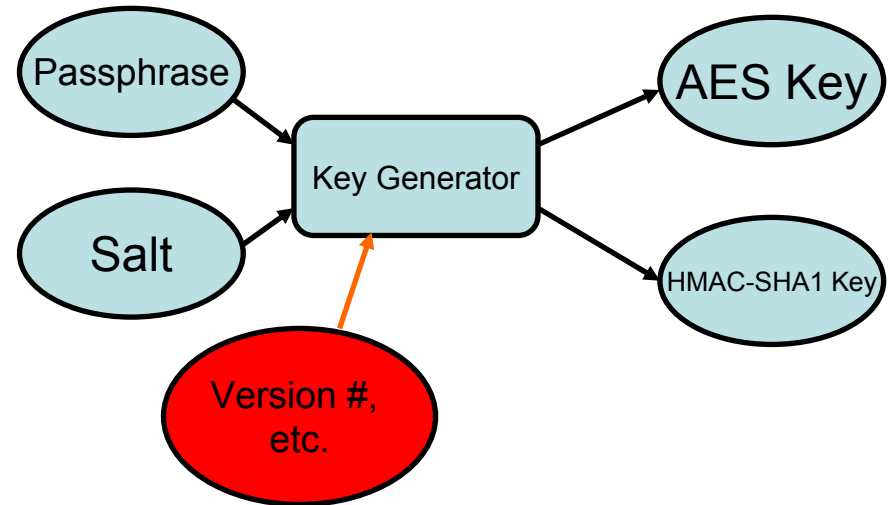
Fixing Compression/Encryption Methods and File Names/Associations

- For WinZip
 - At minimum, the compression type value and file sizes should be MAC'd with the ciphertext
 - Can naturally extend this to include all data necessary to ensure the correct *interpretation* of the data as well (i.e. filenames, dates, sizes, and any other important metadata)

Fixing Information Leakage

- PKWARE's Approach:
 - Provides an option for encrypting metadata
 - Moves data from main file record to central directory record
 - Encrypts central directory (Does not MAC)
 - Archive is no longer parsable
 - Should be a default, not an option
- Encrypt and Authenticate Main/Directory Records
 - Utilize a wrapper file record
 - Encrypted 'original' record file
 - Salt (necessary to derive decryption key)
 - Indication of a wrapper record
 - Inherently solves the past three problems

Fixing AE-2/AE-1 Interaction Problem



- **Problem:**

- Mallory can guess original contents and see if his guess is correct (by changing the version number to correspond to AE-1 and inserting a CRC).

- **Solution:**

- Derive key based on encryption method version number, vendor ID, and encryption strength (in addition to the salt and passphrase).
- Different keys for different versions of the encryption methods.

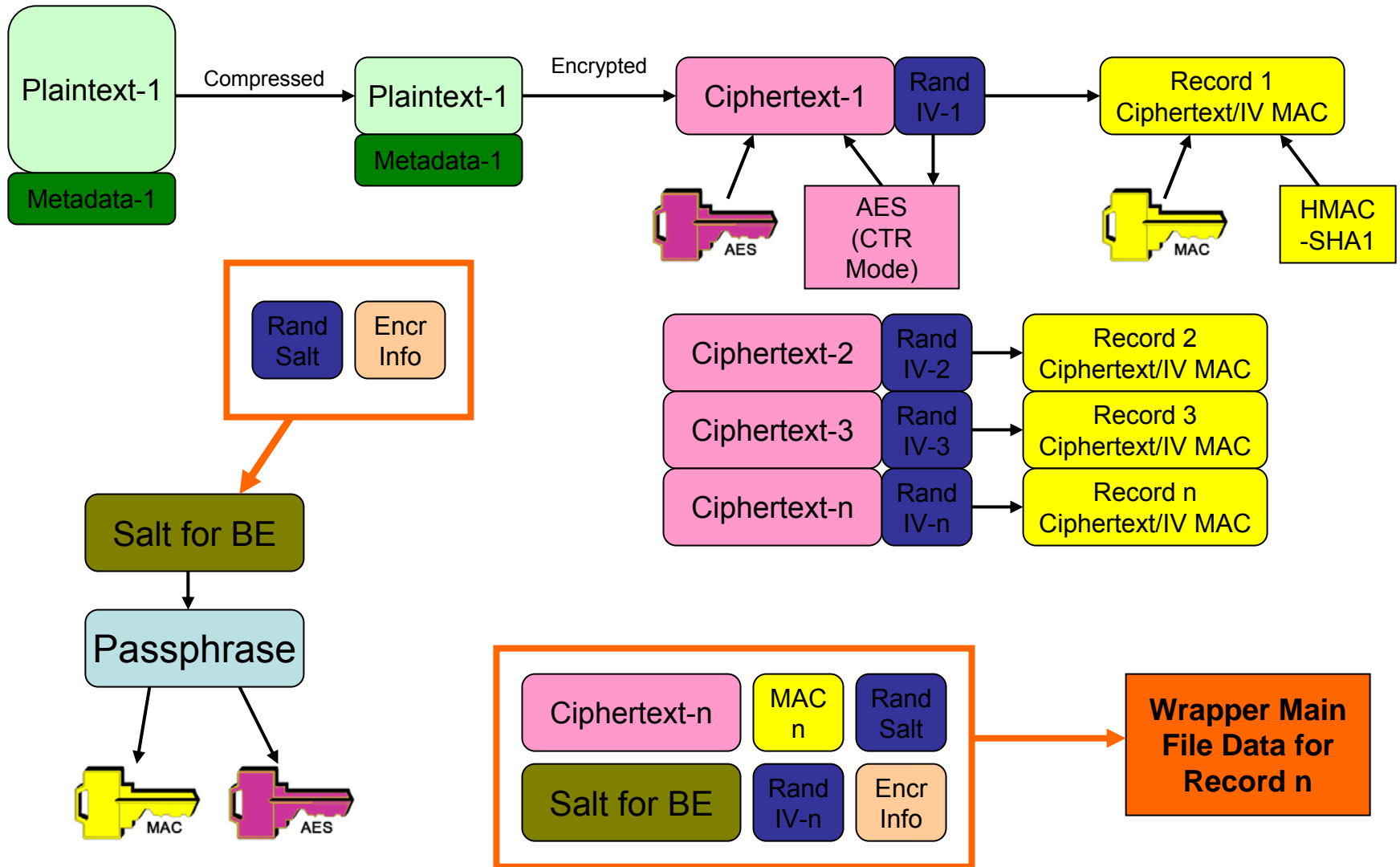
Fixing File Level Attacks

- **Problem:**
 - Mallory can replace encrypted Mallory-Salary.dat file with unencrypted Mallory-Salary.dat file without the recipient detecting.
- **Solution obvious approach:**
 - Authenticate entire archive
 - Might cause problems with efficiently updating large archives (especially archives that span multiple CD volumes)

Fixing File Level Attacks ⁽²⁾

- Authenticate the central directory instead.
 - Include MAC of unencrypted files in central directory so tampering can be detected.
 - Adversary could potentially still turn this off.
- Another potential solution: Just warn the user when an archive contains both encrypted & unencrypted files.

A Possible Instantiation - BE





Actions Taken by WinZip?

- Paper focused on WinZip 9.0
 - Recently, 9.0 SR-1 was released.
 - Fixes a buffer overflow vulnerability.
 - Warns the user when opening an EXE file inside an archive.
 - Doesn't acknowledge correcting any of the issues in this paper.

Conclusion

- WinZip doesn't encrypt or authenticate *all* of the data of interest.
 - *Authenticate everything!*
- Strong cryptography needs to be used correctly, or its useless.