PDF Security Overview: Strengths and Weaknesses

Thomas Merz

PDFlib GmbH
München, Germany
www.pdflib.com

PDF 2001 conference
Scottsdale/Arizona, Nov. 2001

Encrypting PDF Files

- PDF files can be encrypted so that they can only be opened with a password
- Two passwords are used:
  - »user password« for opening the document
  - »master password« for changing the security settings (and the user password)

Security Options

- Security options can be set separately in Acrobat 4:
  - »Printing«
  - »Changing the Document«
  - »Selecting Text and Graphics«
  - »Adding or Changing Annotations and Form Fields«

- Acrobat 5 offers additional fine-grain control over document usage:
  - »Form fill-in and signing«
  - »Document assembly, including insertion, rotation, and deletion of pages and creation of bookmarks and thumbnails«
  - »Allow printing with low resolution only (print as image)«
  - »General editing, comment and form field authoring«
  - »Content access« (for screenreader programs)

How secure are Acrobat's Security Options?

- Files with security options contain everything needed for opening the file – they can't actually be secure!

- The »protection« is rather a plea to Acrobat to disable some features

- Using third-party software instead of Acrobat it's possible to circumvent the security settings:
  - in plain words: it's possible to print »print-protected« files if they do not require a user password!
  - the actual encryption must not be cracked in order to achieve this
  - cracker software does not need to recover the password, but only apply the documented algorithm for opening the protected PDF

How secure is Acrobat 4 Encryption?

- Algorithm is RC4 with 40-bit keys
  - RC4 is used in numerous hardware and software products, e.g. Web browsers
  - 40-bit keys don't provide much security
- U.S. export regulations:
  - encryption products were limited to 40-bit keys for a long time
  - export regulations have been relaxed in early 2000
- Is Acrobat 4 encryption really so weak?
  - www.pwcrack.com/pdf.htm charge US-$ 40 for removing security options and US-$ 500 for encrypted documents
  - www.password-crackers.com/crack/guapdf.html decrypts PDF files in a few days on a single computer (US-$ 29 – US-$ 450)
  - www.elcomsoft.com/apdfpr.html sell software for breaking Acrobat 5 security options and brute-force attacking passwords (US-$ 60)
  - networking multiple computers gives the result »over the weekend«

Improved Encryption in Acrobat 5

- Finally: key length increased to 128 bits
  - 128-bit key can't successfully be attacked by brute force
  - silently implemented in Acrobat 4.05 already (presumably for WebBuy)
  - incompatible with Acrobat 4.0
  - when saving encrypted files you must decide on Acrobat 3/4 (weak encryption) or Acrobat 5 compatibility (strong encryption, more options)
- Encrypt a PDF against a number of known public key certificates:
  - owner needs certificates of all intended document receivers
  - no password required for encrypting or opening the PDF
  - legitimate reader simply uses his private key (profile) to open the file
  - must set up infrastructure for certificate exchange (PKI), or use self-signed certificates
- Cracker programs cannot attack Acrobat 5 files encrypted with 128-bit key as long as documents use an open password

Caution

- Long encryption keys cannot compensate for a poor choice of passwords
- Time to crack 128-bit encrypted PDFs with the ElcomSoft cracker on a PIII/900:
  - four characters: 22 minutes
  - five characters: 9 days
  - six characters: 42 days
  - seven characters: years
- Weak passwords (names, words, etc.) instead of strong ones (which do not occur in a dictionary) result in shorter cracking times

Other Weaknesses of Acrobat Encryption

- Actual segment from an encrypted (!) e-book:

```
256 0 obj
<<
/Type /Encoding
/Differences [ 1 /K /I /N /G /S /T /E /P /H ]
>>
endobj
...
<<
/Type /Encoding
/Differences [ 1 /R /I /D /N /G /T /H /E /B /U /L ]
>>
endobj
```

- Only strings and streams are encrypted in the PDF, but not other object types
- This weakness exists in both Acrobat 4 and 5

Recommendations

- Use digits, punctuation and other special characters in your passwords
- Use passwords with at least 8 characters
- Use 128-bit keys / Acrobat 5, or self-sign security
- Disable font subsetting
- Effective print protection is impossible if you want to distribute your files

E-Book Security: the Goal

- Publish identical digital content to an anonymous audience
  - books, music, software, ...
- Charge individual consumers for content access
- Business model:
  - small fee per copy
  - large number of consumers
  - distribute content at marginal cost via the Internet
- Pirates threaten this business model:
  - purchase a single copy
  - duplicate
  - distribute

E-Book Security: the Approach

- Use a good file format, and distribute the required software for free
- Prevent unauthorized duplication and redistribution
- Individually encrypt each copy, and lock it to the target system
- Users can only consume the content on this particular machine
- Refinement: user can move his license to another machine, but must release the license for the first machine

What's wrong with this Approach?

- We are talking about general-purpose computers
  - no hardware protection available
  - lots of software tools for analyzing and debugging around
- At some point the document must be decrypted in memory
  - fetch the data with your debugger, and save it to disk
- At some point the document must be displayed on screen
  - fetch it from the display memory, and save it to disk
- All »secret« keys and algorithms are implemented in software
  - carefully inspect everything with a debugger, and analyze it
  - write software tools to circumvent the protection scheme
- This is the Internet age
  - when a single hacker cracked the content he can distribute it
  - the hacker can also release his tool, and everybody can break protected files

E-Book Security: Conclusion

- You cannot securely give the content to the consumer, and at the same time hide it
- Protection methods will always be challenged by the current generation of decryption tools
- All protection schemes suffer from this problem, and most have been broken
- In most application areas dedicated hardware is not a feasible solution
- Real-world commerce is based on the scarcity model
  - conventional schemes don't work for digital content which can freely be copied
- If a hacker can break the protection scheme, anybody can
- The whole business model is flawed!

Creatively invent new Business Models

- Funding alternatives from other areas:
  - TV and radio is unencrypted and can be accessed by everybody
  - newspapers and magazines don't profit from sales, but from ads
- Stephen King's approach:
  - publish and charge for individual parts of a novel, and stop publishing when the percentage of paying consumers drops below a threshold
- Charge people for timely access
  - stock information is freely available after 15 minutes
- Accumulate content and switch to a subscription-based service

What does this mean for PDF?

- E-book security is completely different from end-to-end security
- You can't blame ElcomSoft for cracking encrypted PDF e-books
- You can't blame Adobe for an e-book format which has been broken
- PDF is a good vehicle for e-books nevertheless
- Re-consider your business approach with respect to PDF e-books
  - don't try to achieve the impossible
  - adjust your business model to the Internet reality

Resources

- Adobe's e-book pages:
  www.adobe.com/epaper/ebooks/main.html
- Bruce Schneier's talk on »Natural Laws of Digital Content«:
  www.ima.umn.edu/talks/workshops/2-12-16.2001/schneier/DigitalRights.pdf