ROBUSTNESS ANALYSIS OF A PASSIVE PRINTER IDENTIFICATION SCHEME FOR HALFTONE IMAGES

Stephan Escher, Thorsten Strufe

TU Dresden
Chair of Privacy and Security
Germany

ABSTRACT

Printer identification is an important part of printer forensics to protect the reliability of printed documents or to track criminals. While the results of existing identification methods are promising, the robustness of these methods is usually not considered. Therefore, we analyzed an identification scheme which uses the halftone texture of color laser image prints as an intrinsic signature. We evaluated the signature's stability and its vulnerability to forgery. Our results show that the halftone textures depend on different driver settings and can be manipulated. Additionally we show that for specific forensical purposes the identification scheme can be applied to colored text as well.

Index Terms— Digital forensics, Printer identification, Multimedia forensics, Laser printer

1. INTRODUCTION

For centuries, paper and printing techniques have been the foundation for the distribution of information. Today, anybody can print high-quality documents with inexpensive devices as well as edit documents and images easily through high functionality of image editing software. Beside all positiv impacts of this development, printed documents are also used for criminal activities, such as fake IDs, counterfeit money, copyright violation or fake travel tickets. Furthermore, printed documents act often as evidence in a criminal case.

Hence, identifying the source printer, including technology, brand or model, of such documents is an important feature for evaluating their reliability or for tracking criminals. Printer forensical methods provide solutions for this and can be distinguished between active and passive methods [1].

Active methods embed additional information into the document before or at the printing process, e.g. a secure hash of the document or the serial number of the printer. These features, named *extrinsic signatures*, can then be used to identify the printer or to detect forgery. However, these active methods are only useful if the printing process can be controlled, for instance then printing money or IDs.

In the case of printed documents where the printing process is unknown (e.g. ransom notes), passive methods can be used to identify the source. These methods use artifacts in the document caused by the printer. If these artifacts are stable over several iterations and configurations as well as distinguishable among different printers they can be used as identification features, named *intrinsic signatures*, of a specific printer technology, brand, model or the device itself.

Traditional technologies, such as physical [2], chemical [3] or microscopic [4, 5] methods [6], can give good results but are slow, require specialized equipment, educated employees and may destroy the document itself. Alternatively, there exist serveral promising identification approaches which can be performed using common scanning devices and image processing techniques to extract and analyze intrinsic signatures. Such approaches examine e.g. geometrical distortions of text [7] and image [8] documents, the texture and structure of printed characters [9, 10] or color noise of colored image prints [11]. However, these approaches have not sufficiently been tested for their robustness. Possible influences on the stability of the signature, such as type of paper, age of the printer, different toner cartridges or their filling level, driver and corresponding settings and other factors are usually not considered. But the knowledge of the influence of such parameters and the overall signature's robustness is important to judge the practical applicability for printer forensics.

In this paper, we analyze an identification scheme, which uses the halftone process [12, 13, 14]. Its low complexity and overhead makes it especially interesting for broad use – the authors simplify the process to capturing the image with devices as simple as smartphone cameras. The scheme leverages on characteristic halftone textures in *color laser image prints*, and uses these as intrinsic signatures to identify the source printer brand and model. We extend the scheme beyond prints of images to prints that contain solely text, and we evaluate both the signature's stability and resistance to forgery.

[stephan.escher][at]tu-dresden.de

2. BACKGROUND

2.1. Halftoning

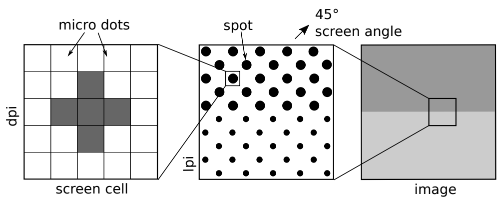


Fig. 1. AM halftoning

Continuous tone images, such as photographic originals, consists of a large range of tonal values. Most common printing technologies only have two operations, to transfer ink or not. Therefore tonal values have to be simulated through breaking up the image into very small dots. This process is called halftoning. The original image is divided into a structured grid with uniform 2D screen cells, where the screen frequency describes the number of cells per inch and is measured in lines per inch (lpi). Each screen cell simulates a tone value of the original image and consists of micro dots which are described by the resolution of the printer, measured in dots per inch (dpi). By selectively filling the micro dots of a screen cell, different color shades can be simulated. The number of possible tone values is denoted by $T = (\frac{dpi}{lpi})^2 + 1$. For multicolor prints, the image is broken into several color channels. For this purpose the subtractive color model is used, which describes the mixture of the primary colors cyan, magenta and yellow to attain a wide range of colors. Typically, a separate black color is included, resulting in a CMYK color model. The halftoning process is applied to each color channel which are then successively printed, overlapping each other.

There exist several approaches describing how to fill the screen cells to simulate the tone values. Overall they can be categorized into Amplitude Modulation, Frequency Modulation and Hybrid Screening [15].

Amplitude Modulation (AM), which is commonly used by laser printers, creates spots that are arranged in regular columns and rows and are equally spaced. Tone values are then achieved by varying the size of each spot in each screen cell. These spots can be created with different shapes such as round, elliptic or square dots.

A problem of such periodic dot placement is the vulnerability to interference effects like moiré pattern, which appears if periodic structures are superimposed. To reduce such interference artifacts the overlapping color screens are arranged in different angles. The best angle between two screens that is least likely to cause such artifacts is 45°. However, in four color process printing all four screens must be angled within a 90° limitation (above the screen would repeat itself). Typical angles are for example 15°,75°,0°,45° for C,M,Y,K.

2.2. Printer Identification Scheme

As explained in section 2.1 by using AM halftoning for color printing the 4 different periodical color screens are overlaid and arranged with different angles related to each other. Ryu et al. [12] proposed that the arrangement of these screens is printer dependent and usable as intrinsic signature for identifying brand and model of a color laser printer. Therefore they developed a method to extract the angles of each color screen. They used 9 color laser printers with 600dpi to print 40 color images each. The printed images were scanned with 2400 dpi and 25 randomly selected sub-regions (128x128px) were extracted and transformed into CMYK color domain. After binarization of each channel by applying Otsus Threshold [16] they used Hough tranformation [17] to estimate the angle values. These values were merged into a histogram. Half of the histograms of a printer were averaged and then used as a reference pattern. The remaining histograms were tested by correlation with the reference pattern. Kim and Lee [13, 14] improved this identification scheme by analyzing the entire halftoning texture of each channel instead of only measuring the angles. Hence the used frequency and spot function are also observed, which enlarges the range of differentiation. They also added the possibility to extract and evaluate the halftone texture with a common camera instead of 2400 dpi scanning. With this approach they reached an average accuracy of 86.14% to identify the printer model of a printed document, out of 5 models from 2 brands [14].

3. EXPERIMENTS AND RESULTS

In our experiments we analyzed the stability of the identification scheme, mentioned above, regarding different parameters and its vulnerability to attackers. Moreover we explored the possibility to apply the scheme on colored text.

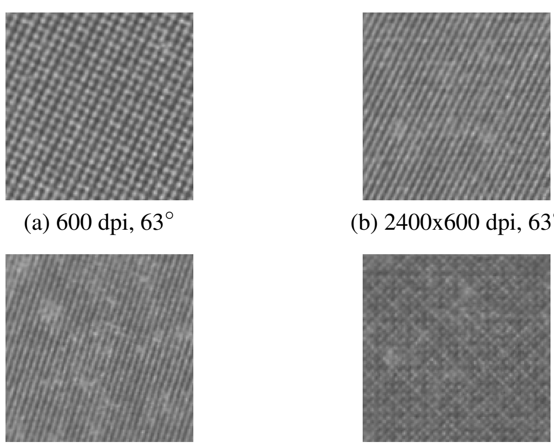
Table 1. List of printers used in our experiments

Label	Brand	Model	# devices
P1	Kyocera	P6021 cdn	1
P2	OKI	MC 361	1
P3	HP	CLJ M553	1
P4	HP	CLJ 5550	2
P5	HP	LJ Pro CM1415fnw	1
P6	Ricoh	MP C3003	1
P7	Ricoh	Aficio MP C305	2

Therfore we created 2 testpatterns T1 and T2. T1 was created for testing the stability and consists of 2 images and 19 squares with different colors (e.g. CMYK colors), to extract the halftone texture as clearly as possible. T2 was prepared to explore the possibility of applying the scheme on colored text. It contains sentences in different fonts (Arial, Times

New Roman) and colors and also some colored icons to compare the image and text halftoning. Overall we examined 206 prints obtained by 9 color laser printers from 4 brands (see Table 1). These were scanned with an Epson Perfection V30 Scanner at 800 dpi resolution.
Stability of the signature was tested regarding different types of paper (recycling vs plain 80g), toners, OSs (Windows 8.1 and Debian 8) and drivers (e.g. Linux PPDs, Universal Drivers, Original Brand Driver) as well as different driver settings.
The color squares of T1 were deskewed and extracted with morphological transformations and contour finding. Out of the squares a subregion (128x128px) was extracted and the halftone texture was analyzed. To estimate the angles of the different patterns we used the Probabilistic Hough Line Transformation [18]. All image processing operations were realized with the help of the OpenCV Framework[1]. Moreover, all texture patterns were evaluated and compared manually.

3.1. Stability of the identification scheme



(a) 600 dpi, 63° (b) 2400x600 dpi, 63°

(c) 1200 dpi, 72° (d) 600 dpi, Dithering: Text, 45°

Fig. 2. Different halftone screens of Ricoh Aficio MP C305 (magenta channel, Win 8.1)

As expected, using different toner cartridges or types of paper did not influence the intrinsic signature. The halftone texture was also stable regarding different printers of the same model (tested with P4 and P7).
However, our results show that the halftone screen is not fixed for a specific printer model. In many cases, the models realize different halftone screens for different driver settings. One of the most influential driver options, measured in our experiment, was the *resolution* setting. Many printer models provide different resolutions, e.g. for fast vs. high quality prints. The impact of the resolution onto the screen angles can be seen in Table 2. By varying the resolution, not only the angle but even the spot function and frequency could be changed (see Fig. 2). These changes were consistent regarding different drivers and operating systems.

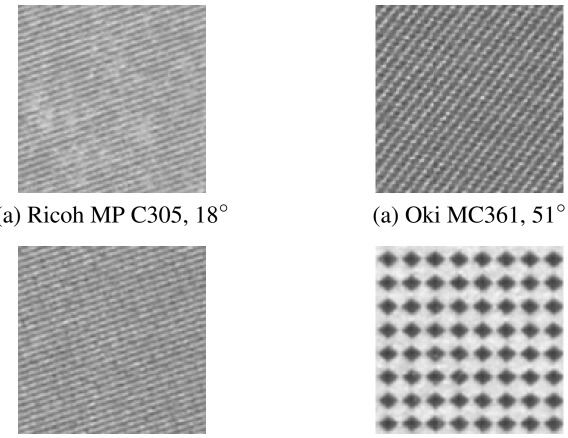
Table 2. Different screen angles through different resolutions

Printer	Resolution	C	M	Y	K
P1	600 dpi	108	72	152	45
P2	600 dpi	75	18	45	45
	1200 dpi	51	128	27	0
	ProQ	104	161	45	45
P3	All	162	16	56	36
P4	All	72	18	0	45
P5	ImageRet3600	75	15	45	50
P6	600 dpi	27	63	0	45
	1200 dpi	18	72	27	45
P7	600 dpi	27	63	0	45
	1200 dpi	18	72	162	45
	2400x600 dpi	27	63	0	117

Additionally e.g. for the Ricoh Printers (P6, P7) there exist driver options named *Dithering* and *Gradiation*. Each of these settings produces different halftoning screens dependent on the resolution. Figure 2 (d) e.g. shows the halftone screen with *text dithering* for P7 at 600 dpi. For both models we get overall 6 different halftone screens through different driver settings. In comparison the HP models (P3, P4) produce the same halftone texture for all tested drivers and settings. P1 supports only 600 dpi resolution and P5 was only tested with a Linux driver that did not support resolution change. This means that a simple change of driver options is mainly sufficient to alter the texture and thereby the intrinsic signature.
Another weakness of the signature is the limited number of useful combinations of screen angles (section 2.1), resulting in possible overlapping of the intrinsic signature between different models. In our testset it can be seen especially within the same brand (P6, P7), but even among different brands (P2, P5), where the spot functions and frequencies were highly similar. Furthermore, the extraction of all 4 halftoning screens is not always possible, e.g. through dark image content. This increases the probability of overlapping (e.g. P2, P4, P5) and negatively affects the identification accuracy.
Based on these results we conclude that the halftone texture is not stable enough to identify the source printer of a document without any additional knowledge. For this purpose it would be necessary to determine every possible halftone texture of a printer model, especially considering different resolution settings. Even assuming that users tend to print with default settings, one should be aware of different halftone textures because of possibly different default settings among different operating systems, likewise mobile devices, and drivers.

[1] http://opencv.org

3.2. Spoofing the intrinsic signature



(a) Ricoh MP C305, 18° (a) Oki MC361, 51°

(c) Spoofed Ricoh screen (d) Spot: Rhomboid, 50lpi, 0°

Fig. 3. Halftone cyan screens of Ricoh Afficio MP C305 and OKI MC361 at 1200dpi and manipulated screens from OKI MC361

The halftone angles and textures measured in our experiments (section 3.1) result from the default configurations of the printer model. But the frequency, angle or spot function are not hardware or device dependent. This means the halftone screen and thereby the intrinsic signature can be manipulated by the user.
For PostScript printers this can be done by sending a postscript file with the corresponding commands directly to the printer. In our experiment we tested this behavior with the *setcolorscreen* command [19]:

```
/sfreq 150 def %150 lpi
/sproc {dup mul exch dup mul add 1 exch sub
} def %round spot function
sfreq 75 /sproc load %75deg cyan screen
sfreq 15 /sproc load %15deg magenta screen
sfreq 0 /sproc load %0deg yellow screen
sfreq 45 /sproc load %45deg black screen
setcolorscreen
```

Figure 3 (a) and (b) show the original halftone screens of the cyan channel of Ricoh MP C305 and OKI MC 361 at 1200 dpi. Fig. 3 (c) shows a manually set screen with a frequency of 170 lpi, an angle of 18°, and a line spot function, printed by the OKI Printer. As it can be seen this setting is a good simulation of the original halftone screen of the Ricoh printer. Furthermore, in Fig. 3(d) the screen is set with 50 lpi, 0°, and a rhomboid spot function printed by the OKI printer.
Additionally these settings can be changed in the specific PPD file. For non PostScript printers there is the possibility of using a Software RIP[2] or digital halftoning to overlap the original screen and make the extraction more difficult.

[2] http://www.rti-rips.com/RIPFeatures/HarlequinPrecisionScreen.html

3.3. Document linking and Application to Text

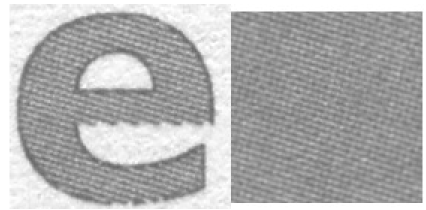


Fig. 4. Halftone screen of a character and an image from HP CLJ M553 (cyan channel, 1200 dpi, 162°, line spot)

Despite this possibility of manipulation, the scheme may still be useful to identify forged documents: simple comparison of expected and detected textures indicates if this a claimed printer has been used, or if the printout is forged - either by altering the printer configuration, or using a different printer. It can also help link printouts to the same originating device, if, for instance, several ransom notes exhibit the same halftone texture. We extended our experiments to prints containing only colored text, to assess the applicability of the scheme for this purpose. Our results yield a match of the halftoning texture in colored text to the characteristic halftoning texture of the devices in 78% of the tested cases (cmp. Fig. 4). Only the Kyocera and Ricoh printers replace the texture with a different dithering in text (cmp. Fig. 2(d)). This dithering, however, is also characteristic and can again be used for identification.

4. CONCLUSION

Our results show that the halftone texture as an intrinsic signature is stable with respect to different types of paper, toner, operating systems and drivers. However, it is highly susceptible to different driver options, especially to the resolution setting. Furthermore, an attacker can spoof or modify the intrinsic signature. For those reasons we conclude that the halftone texture as an intrinsic signature is not suitable to identify a source printer model of a document without any prior knowledge. There are too many possible halftone textures per printer and probable overlaps between different printer models. Nevertheless, it is a promising feature for other forensical applications, like forgery detection or to evaluate the assignment of a set of suspicious documents to one source printer. For such cases the identification scheme can be adapted to colored text. To improve the identification scheme stability and make it more robust against attacks it could be combined with other identification approaches, such as the color noise signature [11].

Overall it can be seen that the analysis of possible influences on intrinsic signatures is important for an accurate evaluation of printer identification schemes and their practical applications.

5. REFERENCES

[1] Pei-Ju Chiang, Nitin Khanna, Aravind K. Mikkilineni, Maria V.O. Segovia, Sungjoo Suh, Jan P. Allebach, George T.-C. Chiu, and Edward J. Delp, "Printer and scanner forensics," *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 72–83, Mar. 2009.

[2] Lukas Gal, Michaela Beloviov, Michal Ceppan, Michal Oravec, and Miroslava Palkov, "Analysis of laser and inkjet prints using spectroscopic methods for forensic identification of questioned documents," in *Symposium on Graphic Arts*, 1993, vol. 10.

[3] Magorzata Szafarska, Renata Wietecha-Posuszny, Micha Woniakiewicz, and Pawe Kocielniak, "Application of capillary electrophoresis to examination of color inkjet printing inks for forensic purposes," *Forensic Science International*, vol. 212, no. 1-3, pp. 78–85, Oct. 2011.

[4] John Oliver and Joyce Chen, "Use of signature analysis to discriminate digital printing technologies," in *NIP & Digital Fabrication Conference*. 2002, vol. 2002, pp. 218–222, Society for Imaging Science and Technology.

[5] Thong Q. Nguyen, Yves Delignon, Lionel Chagas, and Franois Septier, "Printer identification from micrometric scale printing," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2014, pp. 6236–6239, IEEE.

[6] Rashmi Sharma, T. R. Baggi, Amit Chattree, Lav Kesharwani, and A. K. Gupta, "Detection and identification of printer inks-a review report on laser and inkjet printer ink analysis," *International Journal of Current Research and Review*, vol. 5, no. 15, pp. 46, 2013.

[7] Yubao Wu, Xiangwei Kong, Xingang You, and Yiping Guo, "Printer forensics based on page document's geometric distortion.," in *ICIP*, 2009, pp. 2909–2912.

[8] Orhan Bulan, Junwen Mao, and Gaurav Sharma, "Geometric distortion signatures for printer identification," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2009, pp. 1401–1404, IEEE.

[9] A. Ferreira, L. C. Navarro, G. Pinheiro, J. A. dos Santos, and A. Rocha, "Laser printer attribution: Exploring new features and beyond," in *Forensic science international*, 2015, pp. 105–125.

[10] Sara Elkasrawi and Faisal Shafait, "Printer identification using supervised learning for document forgery detection," in *11th IAPR International Workshop on Document Analysis Systems (DAS)*. Apr. 2014, pp. 146–150, IEEE.

[11] Jung-Ho Choi, Hae-Yeoun Lee, and Heung-Kyu Lee, "Color laser printer forensic based on noisy feature and support vector machine classifier," *Multimedia Tools and Applications*, vol. 67, no. 2, pp. 363–382, Nov. 2013.

[12] Seung-Jin Ryu, Hae-Yeoun Lee, Dong-Hyuck Im, Jung-Ho Choi, and Heung-Kyu Lee, "Electrophotographic printer identification by halftone texture analysis," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*. 2010, pp. 1846–1849, IEEE.

[13] Do-Guk Kim and Heung-Kyu Lee, "Color laser printer identification using photographed halftone images," in *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)*. 2014, pp. 795–799, IEEE.

[14] Do-Guk Kim and Heung-Kyu Lee, "Colour laser printer identification using halftone texture fingerprint," *Electronics Letters*, vol. 51, no. 13, pp. 981–983, 2015.

[15] Helmut Kipphan, Ed., *Handbook of print media: technologies and production methods*, Springer, 2001.

[16] N Otsu, "A threshold selection method from gray-scale histogram," *IEEE Transactions on Systems Man and Cybernetics*, 1979.

[17] Sargur N. Srihari and Venugopal Govindaraju, "Analysis of textual images using the hough transform," *IEEE Transactions on Systems Man and Cybernetics*, vol. 2, pp. 141–153, 1989.

[18] J. Matas, C. Galambos, and J.V. Kittler, "Robust detection of lines using the progressive probabilistic hough transform," *CVIU*, vol. 78, no. 1, pp. 119–137, 2000.

[19] Adobe Systems, Ed., *PostScript language reference*, Addison-Wesley, Reading, Mass, 3rd ed edition, 1999.