DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-066

Ambiguity Attack on the Integrity of a Genuine Picture by Producing Another Picture Immune to Generic Digital Forensic Test

Jun Yu; Marvell Semiconductor, Inc; Marlborough, MA, USA
Enping Li; Eastern Kentucky University; Richmond, KY, USA
Scott Craver; Binghamton University; Binghamton, NY, USA

Abstract

Conventional image forgery relies heavily on various digital image processing techniques, which will inevitably introduce artifacts and inconsistency. For the goal of raising suspicion over the integrity of a genuine picture P, we proposed an ambiguity attack not employing any digital image processing techniques.

It works by deliberately producing a second picture P_{amb} containing a target ROI(Region-of-Interest) that highly resembles the ROI in P. Except for the target ROI, the rest of the contents might be dramatically different between P and P_{amb}, so that P_{amb} tells a rather different story from P. Since P_{amb} is not involved with any forgery in digital domain, P_{amb} shall pass generic digital image forensic tests. Furthermore, several measures can be taken to make the ROI in P_{amb} looks more 'original' than its counterpart in P, which induces people to believe P_{amb} is genuine and P is no more than a forgery derived from P_{amb} instead.

The ambiguity created between P and P_{amb} is hard to resolve due to three reasons. Firstly, no digital forensic tool shall identify any artifacts or inconsistency in P_{amb}; secondly, the fact of being able to pass all digital forensic tests still does not assure P is genuine; lastly, determine the chronological order of P and P_{amb} is very hard for general cases.

Introduction

Nowadays, digital multimedia plays a more and more important role in our daily life.

On the other hand, for enormous situations in our world, the truth as to 'what indeed happened in the past' really matters. Quite often, truth can only be revealed by examining an event's record, most likely in the form of digital multimedia. Therefore, Digital Forensics and Anti-forensics, as opponents of an arms race, whichever gains the upper hand would have rights to dictate what is likely to be true.

Take a political scandal picture for example, prove or disprove it might have country-wide impact.

Ambiguity Attack on the Integrity of a Genuine Picture

Unlike most conventional attacks that commit to make a fake image looks genuine, the proposed Ambiguity Attack commits to make a genuine picture looks fake. To be more specific, the goal is to raise suspicion over the integrity of a genuine picture P, by misleading people to believe another deliberately produced picture P_{amb} is genuine and P is no more than a forgery derived from P_{amb} instead.

The steps of an Ambiguity Attack is described below,

1. Since the duplication of the target ROI in P actually takes place in physical world, the attacker should prepare all necessary elements in order to re-create the target ROI and make sure it highly resembles its counterpart in P.
2. Once the target ROI preparation is ready, the attacker should design an environment that is semantically compatible with the target ROI, and set the target ROI into this environment.
3. In order to fine-tune the appearance of the target ROI been captured in picture, the attacker has the freedom to shoot enormous amount of candidate pictures interactively, within a fair amount of time, and possibly use different imaging devices.
4. Among all candidate pictures, the attacker has the freedom to choose the best one as P_{amb}.
5. The attacker then releases this P_{amb} and claims "P_{amb} is genuine, P is a fake forged based on P_{amb}".

A Toy Example

Purely for demonstration purpose, take political scandal picture as an example. Assume someone released a picture P showing that a politician is interacting with another person whom the politician should not have been together with. Been fully aware of its devastating consequence, the dishonest politician then decides to mount an ambiguity attack on P.

Firstly the politician should prepare the same clothes as in P. In the meantime, the politician should choose a different environment and may invite a different person to produce P_{amb}, so that in semantic sense, the scene been recorded in P_{amb} is far from being inappropriate.

In the process of creating P_{amb}, the politician should make efforts to pose the same and reproduce the same facial expression as shown in P, which may not be easy. However, on the other hand, the politician should have plenty of time and a good amount of resources to take enormous amount of trial pictures 'interactively', during which besides fine-tuning his pose and expression, it is also worth to reverse-engineering the relative position and angle of the imaging device at which P was taken. The goal at this stage is to make his appearance highly resemble his appearance in P.

After this painstaking process, the politician chooses the best candidate and release it as P_{amb} with the statement that P_{amb} is genuine but P is no more than a fake based on P_{amb}.

Attacking Philosophy

In order to launch a successful ambiguity attack, the following criteria should be met,

1. P_{amb} shall pass all generic digital image forensic tests.

DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-066

2. The appearance of the target ROI in P_{amb} shall highly resemble the same ROI shown in P. In other words, the target ROI should be replicable by the attacker.
3. Determine the chronological order of P_{amb} and P shall be practically infeasible.
4. Furthermore, it would be better if the target ROI in P_{amb} looks more 'original' than its counterpart in P.

The first criterion always holds due to the nature of P_{amb}. Since P_{amb} is not involved with any digital image processing based forgery, it should be immune to generic digital image forensic tests.

The Freedom of the Attacker

The second criterion may not be easy but meanwhile may not be infeasible in many cases. For conventional attack that relies on digital image processing based duplication, the attacker often has very limited raw material to begin with, and has to struggle with all the inevitable artifacts and inconsistency. As a comparison, in the scenario of the Ambiguity Attack, the attacker has plenty of time and resources in hands.

For instance, if the target ROI in P is a person, the attacker can use image differentiation tool to fine-tune the pose and expression interactively. The attacker can also spend a fair amount of time in reverse engineering the relative position and angle of the camera when P was taken. Moreover, the attacker has the freedom to choose a favorable shooting environment. And as a byproduct of this choice, a different lighting condition from P will obstruct simple differential analysis between (P, P_{amb}). To make it further, attacker is free to produce P_{amb} using different imaging device and/or under different shooting mode such as color tone, dynamic range, ISO etc. from P, which will further obscure direct comparison of the target ROI in (P, P_{amb}).

In a word, this freedom may effectively relieve the second requirement, and makes forensic analysis harder.

Chronological Analysis

There are two types of forensic techniques aiming at determining the chronological order of digital media entities.

Conventional *timeline analysis* relies heavily on various meta data generated by OS(Operating System) and OS File System, as well as auxiliary data such as the *Exif*[1] ancillary tags.[1]

Unfortunately, most meta data is either isolated or separable from the corresponding image data, rather than inherent in the image itself. As a result, those meta data is vulnerable to counterfeit.

The other type of techniques relies on features inherent to an image, hence generally much less vulnerable to counterfeit. The experiment results in [2] strongly indicate that the neutrons in cosmic rays are responsible for most of the hot pixels formed on a CCD or CMOS[2] image sensor.

In [3], this *aging effect* of an image sensor is used to establish the chronological relationship of pictures, with the following working assumptions,

- Pictures under analysis should be taken using the same camera, or more strictly speaking, using the same image sensor.
- The analyst has a set of trusted pictures from this camera with known acquisition time.

As for the proposed Ambiguity Attack, the attacker can purposely break the first assumption by using a different camera to take P_{amb}. Meanwhile, by completely removing all or a large portion of pictures taken prior to P_{amb}, the attacker can either break or largely weaken the second assumption.

In addition, for an image sensor under general conditions, this aging effect is a relatively slow process[2, 4], therefore the analysis granularity been reported in [3] is typically in months rather than in days. This granularity offers the attacker a generous time budget to produce P_{amb}.

Furthermore, since the cosmic ray radiation level essentially determines the hot pixel development rate[2, 4], the attacker can manipulate the hot pixel number by purposely moving a camera to an environment with either low or high radiation level, e.g. a low altitude or a high altitude place repectively.

Cultural Engineering into Play

Cultural Engineering is a concept in *Covert Communication* and is proposed in [5]. While *Steganography* normally commits to hide secret data in statistical sense[6], Cultural Engineering commits to smuggle secret data under the mask of culture.

As a good example, a walkie-talkie iPhone application has been developed in [7]. This application will add a short static noise sound effect to the end of each voice session, mimicking people's experience with a traditional walkie-talkie. The real part is that secret data can be smuggled under the mask of people's cultural experience with a walkie-talkie.

Likewise, in digital image processing, people get used to the impression that a digitally duplicated ROI in a forged image often has inferior image quality when compared with the original. This observation largely reflects the *Data Processing Inequality* in *Information Theory*.[8]

As for the proposed Ambiguity Attack, this cultural experience with forged image may by exploited by the attack to mislead people to believe that P looks more like a fake when compared with P_{amb}.

In order to achieve this goal, the attacker is free to take whatever measures that would produce the target ROI in better quality in P_{amb} than in P. Possible measures include but not limited to, use a better camera, a better lens, a better focus on the target object, a deeper depth of field to cover the ROI better, lower ISO etc.

Once again, as a byproduct, these measures would also obscure direct comparison of the ROI between P_{amb} and P, i.e. make forensic analysis harder.

Practically speaking, thanks to newly developed image sensor technology such as BSI-CMOS[3], camera on smartphone becomes much more popular than traditional digital cameras.[9] Back to the toy example of political scandal picture. If a scandal pictures is more likely to be taken by a smartphone camera, the attacker will have better opportunity to produce a more 'original looking' ROI in P_{amb} when compared to the ROI in P that was taken by a smartphone camera.

[1] Exchangeable image file format

[2] CCD: Charge-Coupled Device. CMOS: Complementary Metal-Oxide Semiconductor.

[3] Backside Illumination CMOS

DOI: 10.2352/ISSN.2470-1173.2016.8.MWSF-066

Discretion in Web 2.0 Era

In the era of *Web 2.0*, as we have observed in many public events, individuals are willing to use their own discretion rather than to follow an arbitration from authority(such as a forensic expert). In many such instances, although their discretion is not as sound as an expert, these discretion may bring widespread impact. This social effect also offers room for the Ambiguity Attack to play.

The Fundamental Limitation of Digital Forensic Test

So far, all attention has been put on P_{amb}. The role of P in the Ambiguity Attack is also worth commenting.

Edsger Dijkstra stated in his Turing Award lecture that for the test-and-fix development methodology used ubiquitously in software development, testing can only show the presence of bugs, *never the absence* of bugs.[10]

The same principle applies to digital forensic test as well — passing all digital image forensic tests still does not assure P is 100% genuine, which helps to consolidate the ambiguity introduced by P_{amb}.

Conclusions

In this paper, we proposed an Ambiguity Attack that raises suspicion over the integrity of a genuine picture P, by deliberately producing another picture P_{amb} that poses an ambiguity question: "which picture is genuine?" Or more crucially, "which picture tells the truth?"

Our attack strategy is quite different from most conventional attacks. While conventional attack commits to make a fake looks genuine, our Ambiguity Attack commits to make a genuine looks fake. While conventional attack makes duplication in digital world, our Ambiguity Attack makes duplication in physical world.

In this attack, since the ROI duplication takes place in physical world rather than in digital world, P_{amb} should elude generic digital image forensic tests.

The established ambiguity is further consolidated due to the fundamental limitation of digital forensic tests, i.e. even if P will pass all digital forensic tests, passing all tests unfortunately does not assure P is genuine.

While good efforts must be spent to make the target ROI in P_{amb} highly resemble the ROI in P, multiple methods can be employed to obstruct differential analysis. Furthermore, there are many ways to make the target ROI in P_{amb} looks more 'original' than that in P, which induces people to believe P_{amb} is genuine and P is no more than a forgery derived from P_{amb} instead.

This attack will be nullified, were it able to determine the chronological order of P_{amb} and P. However, precise Chronological Analysis is by far a very hard problem for general cases.

References

[1] http://forensicswiki.org/wiki/Timeline Analysis.

[2] Theuwissen, Albert JP, and Kleine Schoolstraat. "Influence of terrestrial cosmic rays on the reliability of CCD image sensors." Electron Devices Meeting, 2005. IEDM Technical Digest. IEEE International. IEEE, 2005.

[3] Jessica Fridrich, and Miroslav Goljan. "Determining approximate age of digital images using sensor defects." IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2011.

[4] Chapman, Glenn H., Rohit Thomas, Zahava Koren, and Israel Koren. "Empirical formula for rates of hot pixel defects based on pixel size, sensor area, and ISO." IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2013.

[5] Scott Craver, Enping Li, Jun Yu, and Idris Atakli. "A supraliminal channel in a videoconferencing application." Information Hiding. Springer Berlin Heidelberg, 2008.

[6] Jessica Fridrich. Steganography in digital media: principles, algorithms, and applications. Cambridge University Press, 2009.

[7] Enping Li, and Scott Craver. "A supraliminal channel in a wireless phone application." Proceedings of the 11th ACM workshop on Multimedia and security. ACM, 2009.

[8] Thomas M. Cover, and Joy A. Thomas. Elements of information theory. John Wiley & Sons, 2012.

[9] http://blog.flickr.net/en/2015/12/18/top-cameras-and-brands-on-flickr-in-2015/. (2015).

[10] Edsger W. Dijkstra. "The humble programmer." Communications of the ACM 15.10 (1972): 859-866.

Author Biography

Jun Yu received the B.S. and the M.S. degree in 2001 and 2004 in electrical engineering from Lanzhou University, Lanzhou, Gansu, China, and received the Ph.D. degree in electrical engineering from Binghamton University, Binghamton, NY, USA, in 2011. He is currently a Senior Software Engineer working on security firmware for embedded systems at Marvell Semiconductor, Inc, Marlborough, MA, USA. His research has been concerned with trusted computing on embedded systems, information hiding, and digital forensics. He is a member of the Sigma Xi scientific research society.

Enping Li received the B.S. degree in electrical engineering from North China University of Technology, Beijing, China, in 2002, the M.S. degree in electrical engineering from China University of Petroleum, Beijing, China, in 2006, and the Ph.D. degree in electrical engineering from Binghamton University, Binghamton, NY, USA, in 2012. She is currently an Assistant Professor in the department of Computer Science at Eastern Kentucky University, Richmond, KY, USA. Her research has been concerned with information security, covert communications, and multimedia forensics. She is an associate member of the American Academy of Forensic Sciences.

Scott Craver received both the B.S. degree in computer science and computational mathematics in 1994 and the M.S. degree in computer science in 1995 from Northern Illinois University, Dekalb, IL, USA, and received the Ph.D. degree in 2004 from Princeton University, NJ, USA. He is an Associate Professor in the department of Electrical and Computer Engineering at Binghamton University, Binghamton, NY, USA. His research has been concerned with information hiding, multimedia security, and adversarial issues in signal processing. He was a recipient of the Presidential Early Career Award for Scientists and Engineers (PECASE) for his work in information hiding and watermarking.