The Usage of Geotagging in Hungary

Károly Szommer

Óbuda University, John von Neumann Faculty of Informatics, Hungary
szommer.karoly@nik.uni-obuda.hu

Zoltán Balogh

Corvinus University of Budapest, Institute of Information Technology, Hungary
zoltan.balogh2@uni-corvinus.hu

Abstract: The evolution of ICT technologies and the increasing difference between the technological and social development specified the guideline of the research written in this paper. From the perspective of privacy, one of the most neglected technology is geotagging. Through a survey and technological examinations, this paper describes the dangers from several aspects. These are: examination of file formats, processing images with mobile and PC applications, using offline and cloud applications, measuring user geotagging-consciousness and user behaviour. Dangers were found both on the technological and also on the user side: the technology is more advanced than the awareness of users. The results are pretty surprising, in spite of bad results that were expected. Something must be undertaken urgently to increase the technological awareness of the users.

1 Introduction

The evolution of ICT led to the appearance of new, intelligent communication tools. The level of these technologies rose above the users'. This process led to a rift, and its dangers' are examined in this paper. The geogtagging – what is saving GPS coordinates as metadata - [1] is a very interesting topic. It concerns so many people, but in Hungary only few of them investigate it. A lot of people underestimate the importance of the technology and its dangers up to the present. The technology cannot be called as a new, but by the evolution of the mobile phones and the more intensive spreading of smartphones, its usage is growing continuously. This technology is supported not just by the smartphones, but by the increasing number of webpages too. This technology contributes to our fast-growing digital footprints. [2] Information can be extracted from images created by cameras, mobile phones or any GPS capable devices which can be a huge threat, if gets into wrong hands. [3] This technology implies danger for more and more careless users as it can be automated. [4] There is an endeavour to geotag web contents automatically since

2004. [5] An accurate user profile can be created by using only location information gathered from uploaded images, which can lead to privacy risk. [6] But in spite of this threat, only a very few people know its functions. This is the reason why this topic must be examined in the order of the following steps:

- Technical background of positioning
- Examination of file formats
- Examination of top solutions of various software markets, including both smartphone and PC softwares
- Checking online applications
- Measuring the users' geotagging-consciousness

This paper analyses the possible main sources of geotagging-based privacy violation by testing applications and file formats, and asking people about their geotagging-consciousness. The examination of geotagging as a background technology becomes more and more important because more and more software uses this function of the smartphones without notifying the users. People reveal so many information about themselves by sharing GPS data, mainly because collecting data from the Internet became continually easier by simple crawlers. [7] This is why researchers should pay more attention to privacy-based topics, and this is why this research has been made.

2 The method of research

During the research, at first the users' attitude to geotagging was examined through a survey, then a technical test was made to examine if the threat has other sources. In this paper the technical research is showed first. These two parts are the two main viewpoints of this topic. The existing connection between different types of threats could raise the vulnerability more, if these threats are present. Different methods has to be used to examine different types of threats to give a complete view of the topic.

2.1 The details of technical research

During the software tests, it was important to examine the support of geotagging by modifying pictures, using special functions, sharing them, and even it is important to try if geotags can be removed or not. In all cases Exiftool 8.8.3.0. and Jetphoto Studio were used for checking metadata. These two softwares were used just for maximum compatibility.

The initial file was a geotagged image with jpg extension. In every case the images were converted with ACDSee Pro5 and Photoshop CS5.

The most popular image editors (according to the number of downloads) of mobile OS were tested, in addition the most popular downloadable image editors. The latters were the following: PicsArt, Color Touch, Photoshop Express. Beside the mobile applications the most popular PC software were tested too, and these were Microsoft Picture Manager, XnView, Paint, Photoshop CS5, ACDSee Pro 5 and Gimp.

It was important to examine the online applications – in consideration of the high numbers of their users. Three main groups were created and a geotagged image was uploaded to each of them. Then the image was checked if it still had geographic data in it or not. The examined online applications were the followings:

- Social media: Facebook, Google+, Hi5, iWiW, MySpace, Netlog, Orkut, Twitter
- Image sharing sites: Flickr, Fotolog, Photobucket, Picasa
- Blogs: Tumblr, Blogger (Google), Blogger (Hungarian version), Blogster, Livejournal, Wordpress

2.2 The details of survey

The survey of this research is made by Google Forms. The participants were users who are regarded as "information sharers": the people using social sites a lot. This choice was made because of the importance of this type of users in the view of geotagging dangers. They are the most threatened people, and it can be expected from them to use this technology, these software and smartphones consciously.

The survey consisted of 16 questions, mainly multiple choice and it took about 5-7 minutes to fill. The demographic part was included at the end of the survey, because in the end, the users spend a greater willingness to fill it out. [8] The form measures the consciousness level of people in the field of geotagging, and their knowledge in opportunities in avoiding the threats caused by geotags. At certain technological questions the definition was given before the question.

3 Results / Recommendations

There is a greater threat than it was previously expected. The first type of the threats are the technical specifications of different filetypes and different applications. The second sources are coming from the nature of the "information sharer" users. There is something that must be done to increase the consciousness of the test fillers, both on the technological and sociological side to prevent unwanted data leaks.

3.1 Results of file formats and online applications

The most widespread file formats and online applications were examined during the research. The results of file format examinations were the following:

- JPG: This is the most popular file format, the default output of digital cameras. The smartphones that were used in this research saved the images with this extension. Storing the geoinformation is maximally supported and every software what is able to read metadata can view GPS coordinates in JPG images.
- TIFF: It is similar to JPG. This is because JPG and TIFF are considered the most compatible between various systems. [9] Both file format has an exif header, what is supported by most smartphones, digital cameras and camcorders. Among others, in this exif header contains the geographical positions too. [10] [11]
- GIF and PNG: Compared to the previous formats, these two file formats are not completely supporting geotagging. Not all the softwares can read the geoinformation, not even the Windows Explorer. The similarity between these two extension is because of PNG was intended to be the long run replacement of GIF. [12]
- BMP, ICO, IFF, JPG2, PCX, RAS, RGB, RSB, SGI, TGA, WBMP: These formats cannot contain GPS coordinates.
- RAW, PSD, EPS: It depends on the given software that these formats can contain GPS metadata or not.
- It is important to mention that the most popular file formats can contain the geoinformation. These are so widespread, that smartphones users are using them almost all the time. That's why more detailed research is needed in the field of applications.

Geotags are saved when the photo is taken. There are many applications in the markets of every smartphone platforms that are capable of editing images, but only the most popular and the built-in native applications were tested. The smaller corrections are made on the smartphones most of the time, these applications have to handle geotagging right to support users' consciousness.

As Table 3-1. shows, the three most popular native mobile OS image editors keep the geoinformation. The problem is that they do not let their users know about it, so the sensitive data remains in the images.

Function	Metadata persists
sending e-mail	yes
compressed sending	yes
cropping	yes
turning	yes
applying special effects	no
correcting	no

Table 1
Geotagging metadata persistence at native softwares of smartphones

After the analysis of the results of native mobile applications, it was expected that the most popular mobile image processing software will handle this type of metadata a better way, but the results caused disappointment. None of them supported the work with GPS metadata at all. This type of information disappeared from the photo. The most popular software was Photoshop.

In contrast to the mobile applications, softwares for PC fully supported GPS metadata. (Table 3-2.) The most surprising result came from the examination of Paint. It seemed to be the most simple image editor but it has full support.

Program	crop	turn	mirror	special effect	correction	drawing	text	contrast
Paint	x	x	x	-	-	x	x	-
MS Picture Manager	x	x	x	-	x	-	-	x
XnView	x	x	x	-	-	-	-	x
ACDSee Pro 5	x	x	x	x	x	x	x	x
Photoshop CS5	x	x	x	x	x	x	x	x
GIMP	x	x	x	x	x	x	x	x

(Signs and abbrevations used: x persists, - doesn't have this function)

Table 2

Results of the PC image processing softwares

The market share of online applications are increasing rapidly. In consideration of the values of personal data, it was relevant to find out which sites store and show the metadata in the image available on them.

At first the social sites were in the center of the research. One of the most popular social sites is Facebook. [13] In 2011 it has nearly 750 million users. [14] Later in 2014 it grew to 1.3 billion. [15] 78% of Americans use Facebook on mobile phones even from the registration. [16] This data shows us the importance of the connection between image metadata and smartphones. On Facebook the GPS data can be given voluntarily, and even superficially. The uploaded images don't contain this type of information.

Whereas the Google+ save these data and with a simple exif reader anybody can read them if he is able to see the images, which is a huge threat. The other social media sites (Hi5, MySpace, Netlog, Orkut, Twitter) don't support the displaying of these metainformation.

Not just the social media sites but the image sharing sites can have the ability to show geoinformation. The only site where metadata was deleted is Fotolog, and it doesn't support them at all. (Table 3.)

Site name	GPS metadata can be shown
Flickr	yes
Fotolog	no
Photobucket	yes
Picasa	yes

Table 3
Image sharing sites and geotagging

In this category we found the first online page that definitely tried to increase geotagging consciousness. It was Flickr. The page stored all GPS metadata but nobody could get it unless the approval of the user. This page draws user's attention to threats caused by geotagging.

On the other hand Photobucket and Picasa have no protection. The sensitive data on these page can be read very easily. It is important to mention that Picasa is the second service of Google that fails on the security test. Who thought that Google, who declared on 1st March, 2012 [17], that its pages are more secure, don't pay attention to such threat.

Only one of the examined blogs were unable to support geotags, which was Blogger, the only Hungarian one. Most of the time people writing blogs want to stay anonymous. With self-taken images, this is a security risk because with geoinformation users can be identified.

As it can be seen, in this case the data vulnerability comes from different layers of the image-sharing process: image capturing, file specifications, image editing processes, sharing on different websites. In this case the threat coming from technical side could be eliminated by warning the users each time a data leak could occur because of metadata sharing. This could be done on different layers: before the first start of the camera, or before the first upload of a geotagged image.

3.2 Survey

The form was filled out by 524 people, they were mostly smartphone users. The fillers were mostly females (64.3%). The distribution of the ages can be seen on Fig. 1.

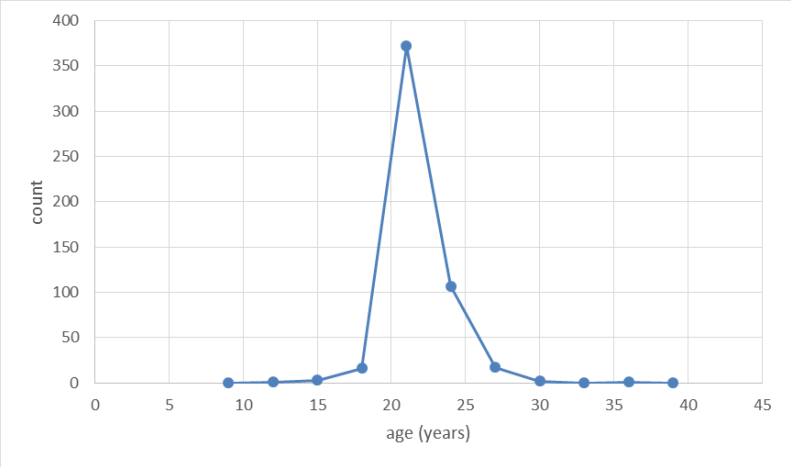

Figure 1
The distribution of ages of the answering people

Most of the answerers live in Budapest, in Pest county or in a bigger city. (Fig. 2.) Most of them (66.2%) didn't know what geotagging means, or what it can be used. After they get to know this, 44.6% declared that it doesn't comes to their mind that the metadata can be viewed with just a little computer knowledge. 63% of them thought that this function cannot be switched off.

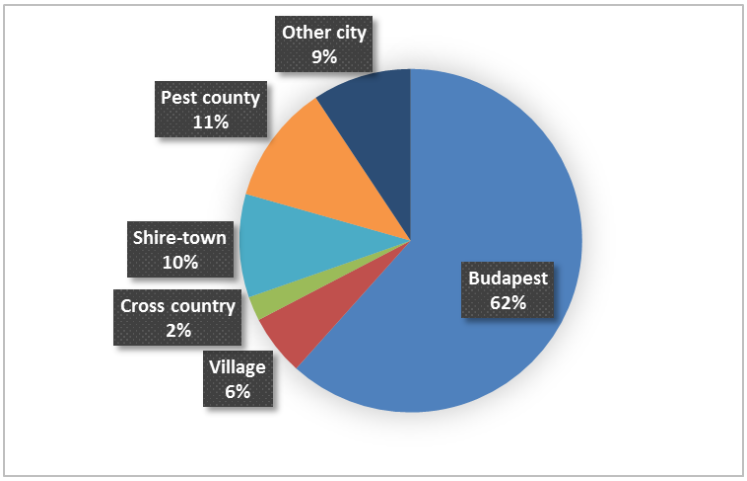

Figure 2
The distribution of inhabitation

On Fig. 3., the correctness of classification is shown of the mobile phone that are capable of geotagging. 48% of the answerers could not determine if their phone is able to use GPS metadata. In the worst condition were the people who answered that they definitely know that their phone is unable to use geotagging (6%). These people are in increased danger while posting photos, because they live in the false illusion of security.

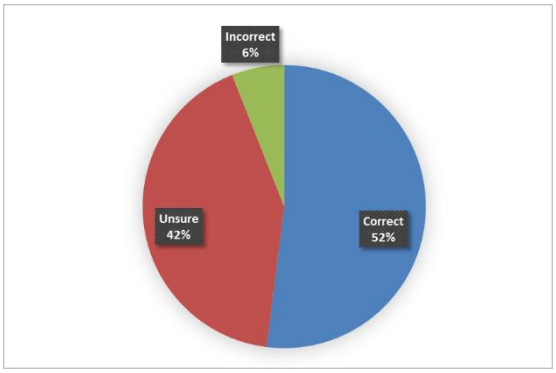

Figure 3
The rightness of the answers if the phone is capable of geotagging or not

In consideration of geotagging, as an IT-oriented thing, the younger (<15) and the older (>30) age-groups have lower consciousness. This is why the results of a survey participating from the whole Hungarian population would be worse.

People sharing self-taken pictures are in increasing danger, because they mostly use the same file formats that are totally support geotagging. Companies are trying to maximize their profit by using new technologies as they can attract more people to themselves with them. User safety is secondary for most of them. [18] Through more and more online applications are flowing the metadata to the open internet from the more and more popular smartphones.

The second source of the threat examined in this paper comes from the sociological part. It seems that geotagging and conscious information sharing is not the part of today's general technology training. People mostly trust the manufacturers, the huge companies. They think that these companies will keep their data in safety just because of so many people use their services.

4 Conclusions

The main sources of geotagging-based privacy violation were identified and the specific deficiencies were identified. As it can be seen, the sum of these problems raises the dangers to a higher level because of the many aspects of vulnerability: not just the people -using geotag-based services- are careless, but the companies making those services too. The possibility to secure these type of services are given, but only a few companies make use of this solution.

References

[1] Valli, Craig – Peter, Hannay (2010): Geotagging Where Cyberspace Comes to Your Place, Security and Management, pp.627-632.

[2] Éva, Túri (2011): Digitális lábnyom keletkezése és kezelése

[3] Weaver, Stephen D. – Mark, Gahegan (2007): CONSTRUCTING, VISUALIZING, AND ANALYZING A DIGITAL FOOTPRINT, Geographical Review, 97(3.), pp.324-350.

[4] Martha, Larson – Mohammad, Soleymani –Pavel, Serdyukov (2011): Automatic Tagging and Geotagging in Video Collections and Communities, Proceedings of the 1st ACM international conference on multimedia retrieval, pp.51-58.

[5] Einat, Amitay – Nadav, Har'El – Ron, Sivan – Aya, Soffer (2004): Web-a-Where: Geotagging Web Content, Proceedings of the 27th annual

international ACM SIGIR conference on Research and development in information retrieval, pp. 273-280.

[6] Julien, Freudiger – Shokri, Reza – Hubaux, Jean-Pierre (2012): Evaluating the privacy risk of location-based services, Financial Cryptography and Data Security, pp.31-46.

[7] Károly, Szommer (2012): Webes adatbányászat, Acta Carolus Robertus, 3.(1.), 2012. pp.137-142.

[8] Edit, Bódi - Tamás, Gyulavári - Ágnes, Neulinger - Ariel, Mitev - Judit, Simon - Krisztián Szűcs (2011): A marketingkutatás alapjai

[9] W. Fulton (2010): Image File Formats – JPG, TIF, PNG, GIF, at: http://www.scantips.com/basics09.html

[10] T., Tachibanaya (1999): Exif file format, at: http://www.media.mit.edu/pia/Research/deepview/exif.html

[11] JALOXA (2011): EXIF header, at: http://www.jaloxa.eu/webhdr/exif.shtml

[12] Before IT's News (2011): Six Popular Image File Formats: What Are They and How Are They Best Used? – A SPN Exclusive Article, at: http://beforeitsnews.com/business/2011/07/six-popular-image-file-formats-what-are-they-and-how-are-they-best-used-a-spn-exclusive-article-804116.html

[13] Facebook (2014): Facebook info, at: https://www.facebook.com/facebook/info

[14] Alfonso, Serrano (2011): The Social Media Explosion: By the Numbers, The Fiscal Times, at: http://www.thefiscaltimes.com/Articles/2011/09/12/The-Social-Media-Explosion-By-the-Numbers.aspx#page1

[15] Statistic Brain (2014): Facebook Statistics, at: http://www.statisticbrain.com/facebook-statistics/

[16] Josh, Constine (2013): Facebook Reveals 78% Of US Users Are Mobile As It Starts Sharing User Counts By Country, at: http://techcrunch.com/2013/08/13/facebook-mobile-user-count/

[17] IT Security Training Australia (2013): Google's Privacy Policy Changes Come into Effect March 1, 2012, at http://www.itsecuritytraining.com.au/articles/google%E2%80%99s-privacy-policy-changes-come-effect-march-1-2012

[18] Puskás Tivadar Közalapítvány (2012): Biztonságosinternet Hotline, at: http://www.biztonsagosinternet.hu/