Digital Footprint: Pros, Cons, and Future

Prof. Anjana R. Arakerimath[1], Pramod Kumar Gupta[2]

MCA Department. Pimpri Chinchwad College of Engineering, Savitribai Phule Pune University
Pune, Maharashtra 411044, India.

Abstract- **We live in the world of digital era of social networking and online information exchange. There are different online services used by users. Digital footprint help web service provider to target appropriate user for their web content and services. In this paper we are going to see what is digital footprint is and its classifications. What are privacy issues it leads to? Role of social networking site in context of generating digital footprint.**

Keywords- ***Footprint, Digital Footprint, Privacy issues, Passive footprint, Active footprint, Social networking objects.***

I. INTRODUCTION

A **digital footprint** is the data that is left behind by users on digital services [1]. In simple words, it is the data about the data that user is searching or using online.

Today almost one third of the entire population is using the internet. We can say that almost everyone is on the internet or soon will be. Today, the internet is not only limited to web surfing or sending email it has taken completely new meaning in which social networking play an important role. Social networking sites like Facebook, Twitter, Google+ and LinkedIn are dominating the entire web. People are sharing almost everything they can on the internet like how they are feeling, what they are doing, their hobbies, what they like, what they don't like etc. Everything is posted online. Since these activities are posted online hence they must be kept somewhere and by somebody. The somewhere may be the private servers, grid computing or the cloud computing. And somebody can be the companies like Facebook, Google or some other companies on which we are doing online activity. Whatever the information is stored from your online activity is nothing but your online activity log or we can say your online digital foot print. A digital foot print can be anything that you share directly or indirectly online whether it is a simple text, image, audio, video etc. anything which is traceable back to you. In Fig. 1 we can see different ways through digital footprint comes into life and how it is processed and utilized.

It can be evolved by contents you have posted online like comments, likes, image shares etc. It can also be evolved by the contents that your friends have posted about you like image tags, hash tags of your name etc.

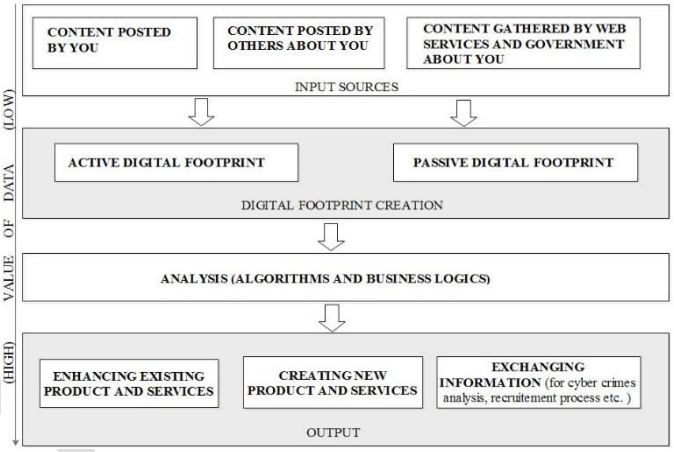

Fig. 1 Digital Footprint Creation.

Sometimes digital footprint are also generated because of some web services which gathers information about you like your IP address, web browser and location information etc.

A. Classifications

The digital foot is further classified into two categories which are passive digital footprint and active digital footprint [1].

1) Passive Digital Footprint: Passive digital foot print is created when data is collected about some online activity without client activation [2]. Means data is collected implicitly and most of the time user is unaware of such data collection.
For example, in an online environment whenever a user browses any website then the websites can trace his geographical location through user's IP address.
In an offline environment, footprints can be stored by system in files or system activity log files through which admin of that system can see action performed by the machine, without being able to see who performed them. The owner of the system can put some software which keeps the activity log of user for a particular period of the time like in weekends.

2) Active Digital Footprint: Active digital footprints are created when a user willingly releases personal data for the purpose of sharing information about him. Here data is provided explicitly and user is aware of such data collection by the companies [2]. For example, In an online environment when a user creates a social networking

profile or comment on some post or article then in such scenarios user is creating active digital footprint of himself.

In April 2010, Facebook has launched its "Open Graph" API. It is a concept through which developer make social object on Facebook platform [3]. Social object is everything that you see on its social networking sites. A brand like Apple Inc. or simple product like bath daily bath soap Dove, everything is social object which is connected with the 'like' clicks of millions or billions users all over the world connected through Facebook. Therefore, whenever you like content on Facebook you liking a social object which is a real life entity. And you can also create social object of any real life entity you want, it doesn't matter the entity is real life place where you live, a song, a college, an organization, a birthday event or simple product or post that you want to promote.

Through 'open graph' Facebook is trying to connect each and every real life entity online. And these online connections of social object has created tremendous amount of digital footprint and these footprint is not only limited to users but its extended to products or events or whatever social object has been connected with user. Here mostly the digital footprint takes place with help of user's active footprint. For example when a user like product called iPhone then he is generating active digital footprint which tells that this user like this particular product but it's also generating passive digital footprint in the background which user may be not aware of and that passive digital footprint may contain information like user geographical location, age, gender, profession details which in turn generate a new digital footprint with combined footprint of the user of same genre and this newly generated passive digital footprint can tell us people of what age group or profession are liking or using the iPhones.

Digital footprint becoming more complex as the number online users and their connection with each other and social content are increasing day by day. For example consider a Facebook user from USA and a Facebook user from Australia who doesn't know each other and there is connection between them not even any common Facebook mutual friends can be indirectly connected to passive digital footprint based on their active footprint generated by them though the clicking of LIKE button on the same social object. The object can be a brand or Article or blog page.

II. Related Work

Digital footprint is one of the important concerns in the era of social networking and online information sharing but its origin started since the World Wide Web has been developed. There are few research has been done earlier.

In 2007, Tony Fish pointed out the concept and the possible danger of digital footprints in a self-published book [2]. It explains how closed loop takes data from the open loop and provides this as a new data input. This new data determines what the user has reacted to, or how they have been influenced. The feedback then builds a digital footprint based on social data, and the controller of the social digital footprint data can determine how and why people purchase and behave.

It also covers one of the important concern in digital footprint world that is, is privacy a two-sided digital business model where your privacy will be someone else's business?

On October 2012, there was research a paper published on the topic "Private traits and attributes are predictable from digital records of human behavior" [4]. It explains how the private and traits and attributes of an individual can be predicted through its online shared and exchanged information i.e. through its digital records.

III. Pros

Although digital footprint may sound one of the negative aspect of online activity as log of almost every activity is maintained through active or passive digital footprint. But there so many advantage of digital foot print which helps in making our online experience more friendly and accurate. Without digital footprint our online world may not be very easy like it is today.

Google Ad sense uses our digital footprint to serve us appropriate advertisements. It uses our passive digital footprint but with our permission as at the time of sign up they mention it in their terms and conditions that they will use our web surfing data to serve us more appropriate advertisement based on the content we are searching for. If the Google have not used our digital footprint data then it may be happened that we are searching content related to job and we are seeing advertisement of international holiday packages, which make no sense to users and it will be also not beneficial for the company of which the ad has been shown as user will never going to visit that ad. Hence with the help of digital footprint Google target the appropriate user for showing the ad which is beneficial for the both parties the user and the companies of who's ad has been shown.

YouTube is also one of the product provided by the Google and it also uses our digital footprint to improves our related vides list. For examples if we search for the movie called inception then in the 'suggested video list' it show us the movie related with same genre or the same cast of inception.

The Facebook also uses or digital footprint to give us better social networking experience. It display the advertisement based on our likes and friends likes. It displays suggested friend list based on our friend list and mutual friend list and there some other algorithms based on our likes of pages etc.

Now a days we have GPS enabled mobiles phones and tablets which add our geographical information to our digital footprint. It also helps us to serve better web services like Google Maps helps to find locations more

accurately with help of our past search data and current GPS location.

IV. Cons

Digital footprint help us to serve us better work experience but one the disadvantages of digital footprint is that most of the time data is collected about client online activity without client's awareness specially in passive digital footprint.

Today we can browse so many webpages within a minute and many more in hours. We visit different types of website with different terms and conditions which are very long to read and most of time we just agree with terms and conditions without even reading it and start using the website and services which results in giving legal rights to website to collect our online activity information. According Microsoft's 2013 Privacy Survey Results [5], Majority of people don't bother about terms and conditions of websites. The survey result is shown in Fig 2.

We can control our digital footprint if we do proper privacy setting and take precautions while surfing, but today our digital footprint has become very large to remember and maintain.

Digital footprint can help others to predict our private traits which we don't want share online. Based on our Facebook likes, comment on different online blogs, online product purchases list etc. a behavior or consumer analyst can tells us our product taste and preference. If you are a regular web surfer and uses online social networking sites regularly to exchange information then if you search your full Name on the Google

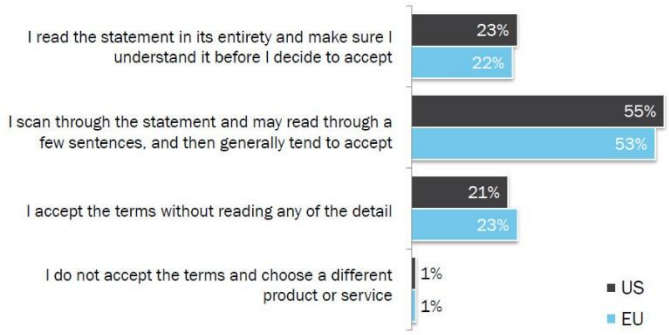

Fig. 2: Terms of Privacy Behavior

you might surprise the result shown by the Google search, you will be more surprised if you search your name in image search tab. It might show your picture that you have putted long time ago on some social networking site and forgotten it.

The social sites gives your privilege to control your online sharing things but sometimes users don't even understand what is happening

For example, if we tag example of image tagging feature provided by the Facebook. Image tagging is a feature which helps you to tag or put your name on image or picture if you are present on that image. It helps you to identity particular person in a group of picture. But sometimes, most of user uploads any picture or wallpaper and they tag all of their friends. Hence even if you are not in that picture you will be become part of that picture as by definition of tagging provided by the Facebook we tag whatever whoever present in that image.

That is the reason when we search our name on Google we get surprised by seeing the picture which doesn't belongs to us and become part of wrong digital footprint. Such small things play very important role in user privacy.

The social networking websites like Facebook, Google+ gives you control over what you share online through their privacy setting features and in terms and conditions they also mention that directly or indirectly the overall rights of the content you share will belongs to the these social networking sites only. You may deactivate your account but there is no guarantee whatever you shared online before deactivating your account will be destroyed permanently after deactivating the account. As most of the social networking sites never destroy their user information from their server, it may not appear on the web page of social networking sites but they keep data of every single user, even if the user has left the site, they use data of those users to analysis like what kind of users have left the website, how long they have used website, before leaving what were there last activity etc.

It's not only the teenagers and adults who are in the web of digital footprint but new born baby is also becoming part of digital footprint even before they are born. In 2010, Internet security company AVG surveyed [6] mothers in North America (USA and Canada), UK, France, Germany, Italy and Spain, Australia, New Zealand and Japan, and found that 23 percent of children have had their pre-birth scans uploaded to the Internet by their parent – establishing a digital footprint even before birth. Fig. 3 [7] is based on AVG survey result which shows key results regarding digital footprint for new born babies.

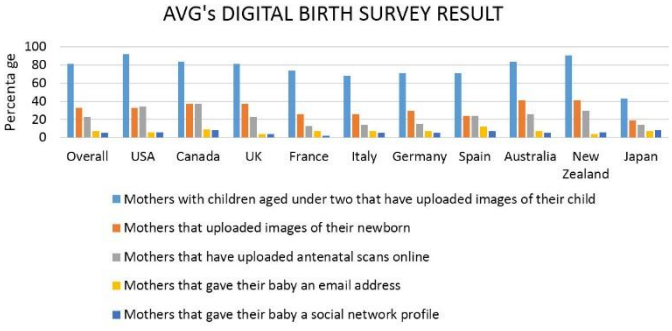

Fig. 3: AVG Digital Diaries Campaign Survey Report.

Today there are some people who spend more time online then in their real life. Some people make fake online social networking profiles and email IDs. They put wrong information on their online profile which is completely different from there real life personality which leads to

generation of improper digital footprint. And these wrong digital footprints gives wrong analytical data to the web services which more rely on active digital footprint. That is the reason most of the web services prefer passive digital footprint which collects data as per there algorithms and logics they developed for collecting data from the user.

Even if the person dies in real life, his digital footprint will be live forever on internet. One of the important concern about our digital footprint is that who has right over the digital footprint of the user, the company who is collecting it or the user who is creating active or digital footprint through its online activity. As digital foot print provides so many advantages to companies like in understanding customer preference of products, analyzing his changing behavior etc. Digital footprint may help customer to understand there customer or user better. But sometimes it's violating the user online privacy as sometimes data is collected without knowing of user.

So who should be responsible for responsible for protecting person's online privacy? According to Microsoft's '2013 Privacy Survey Results' [5], protecting consumer's online privacy is a shared responsibility which consist of Individuals, Companies and Government.

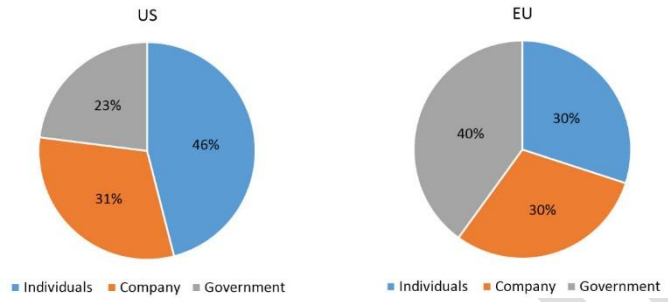

Fig. 4: Responsibility for Protecting Consumers' Online Privacy

V. MANGING DIGITAL FOOTPRINT

TABLE I
WEBSITES FOR MANAGING DIGITAL FOOTPRINTS.

Reputation.com	It is an online reputation management. The company's products and services are designed to help businesses and individuals monitor, manage, and improve the way they appear online, across search engines, online reviews, and social media.
Unroll.me	It is the easiest way to manage your inbox. Unsubscribe from unwanted email subscriptions.
Spokeo.com	It utilizes deep web crawlers to aggregate data. Searches can be made for a name, email, phone number, username or address. In response to privacy complaints and lawsuits, the site allowed users to remove information about themselves through an "opt-out" process.
intelius.com	It provides information services, including background checks and identity theft protection, and post transaction marketing.
teensafe.com	It allows parents to monitors children's text, calls, social media, phone and more.

Although now a days it is almost impossible to erase our digital footprint as we just keep using new web services and after few years we don't even remember number of registration, subscription or other online activities that we have done. But there are websites through which we can see our online digital footprint, our online projection. Table 1. contains few web services which can be useful for users to monitor themselves as well as their children also.

VI. FUTURE

The future of digital footprint can be very complex and wide. If we think of its positive side, with the rapidly increasing number of online content and online user's digital footprint can play very crucial role in co-relating the content of the web and serving it to the user. There are few things that we have predicted based on our analysis and rapid changes happening in technology. These things might be come into to life with in a decade or maximum decade and half.

Today in recruitment process of Multinational companies, the recruitment team also have digital footprint analyst. They goes through digital footprint of candidate available online. So in future may be your online digital footprint becomes one of important attribute in recruitment process.

Today we have many matrimonial and dating websites which finds the perfect matches on the basis of person's likes and dislikes, In future there will be websites which will analyze the person's digital footprint to make instant perfect matches which may be more effective.

In future there will be many companies who will do business by providing services like improving person's digital footprint or destroying person digital footprint. There may be services which will take responsibility for managing person digital footprint from birth to death.

We cannot judge or know about any person without spending time with him or her, but in future there will be completely different social network or services which will focus only on person digital footprint, and it will display all the attribute of person in the form of levels like confidence, stress, IQ, integrity, stability etc.

Our perception about an individual differs from person to person, but in future there might be some generic perceptions categories which will be assigned based on persons digital footprint on the web.

Today we have money system, online payment systems, even we have bitcoins. In future may be a person's digital footprint becomes one the aspect of money system for using or exchanging services. For example, In future there might be some things or services and for buying or using such things or services only money will be not enough, your digital footprint should be up to required level.

VII. Conclusion

There may be few disadvantages of digital footprint but in future it is going to play very crucial role in serving right content to web users and it will help online web service providers to understand their clients and serve accordingly. Every second some active or passive digital footprint generated online and it will keep generating as long as someone is using the some online content. As it helps in providing related web content to its appropriate user.

In the end it's on the user how they manage their digital footprint and control their privacy issues.

References

[1]. Digital footprint [Online]. Available: http://en.wikipedia.org/wiki/Digital_footprint.

[2]. Tony Fish, *My Digital Footprint: A two-sided digital business model where your privacy will be someone else's business!*

[3]. Facebook Open Graph API [Online]. Available: https://developers.facebook.com/docs/graph-api/overview, http://ogp.me/

[4]. Michal Kosinski[a,1], David Stillwella, and Thore Graepelb, "Private traits and attributes are predictable from digital records of human behavior"

[5]. 2013 Privacy Survey Results [Online]. Available: http://www.microsoft.com/en-us/twc/privacy/data-privacy-day.aspx

[6]. AVG Digital Diaries Campaign Survey Report [Online]. Available: http://www.avg.com/digitaldiaries/2010

[7]. Digital Birth: Welcome to the Online World [Online]. Available: http://www.businesswire.com/news/home/20101006006722/en/Digital-Birth-Online-World#.Ve_2ohGeDGd