

The Socially Acceptable Social Networks

A Reality Check on Information Engineering for Internet

Sumith Kumar Puri

Abstract—The world of internet is a myriad of information. This information can be classified under multiple categories. Most of this information is not rightfully branched. The earlier generation of information access such as a library had no such qualms. Though there are multiple solutions to such issues on the internet, most of them are a cure for an otherwise avertable situation. This article is related to the prevention of such information patterns of data on the internet, especially across social networks. It also discusses a way to safely bifurcate the internet in terms of content and relevance. This article is also an effort to encourage and create a thought process to provide a more relevant cyber society for the current and future generations. It should be able to direct social network creators and cyber law enforcers in building more socially acceptable networks.

Index Terms— Information Engineering, Cyber Society, Cyber Policing, Internet Engineering, Cyber Laws, Social Networking



1 INTRODUCTION

Picture this; you were to search across the internet for your long lost school friend – with whom you have no contact directly or indirectly. The most common option would be to first fire up your favorite social network and search through a long list of similarly named people. Some people may also use person directories, public contact lists or simply search engines. But during this search you would obtain high volume of other irrelevant and possibly unacceptable content. This information usually includes escorts, adult companions, marketing, pornography, malware, email scam links or even gambling sites. This is only a subset of the other high amount of violent, racist, adult, fraudulent and scam links that may appear in your search. They will not only evoke negative feelings but may even take away your vital information like credit card or passwords.

Now, apply the same to your teenage cousin searching for his primary school friend!

2 DETAILED DESCRIPTION

Internet, though still in its commercial or public infancy in terms of relative age with other important modes of information has come a long way in enforcement. But we still are far from what should be the actual norms itself. The World Wide Web provides us with almost any information in the universe in a flash. Most of the times, this information is not legally, ethically or socially solicited. This creates a plethora of issues for a user. Let's first understand a more holistic view of this problem.

Easier access to sensitive, adult or racist information especially to children is capable of creating a society which is violent, intolerant and insecure. Also, wrongly placed information at various sources decreases the efficiency of the net user in general. It not only creates a barrier in obtaining the right information but also information chaos. In addition to this, there is also abundance of irrelevant or profane terrorist, religious and advertising information. It is also

common on the internet for a user to register for one of the e-magazines, e-stores or mobile texting sites. This is mostly done using the most frequently used email id or even cell-phone numbers. As you would have already experienced your inbox then becomes an entry point for multiple internet scams, chain mails, malware or adware. Apart from this, since there is a high amount of freethinking on the internet there is an issue with irrelevant naming, redirection, content when the actual context differs. This becomes a greater issue with multiple sub domains, free listing on search engines, ad driven linking, pay for click or pay per visit web sites. Though very rarely, you might have heard of internet bullying or even unsolicited meeting mails which could have led to financial or physical distress. As the extent of internet grows, it will lead to more such cases and ones which may prove even more threatening. Also, once an internet user is able to identify the information as fraudulent or misleading; he or she is unable to usually find a quick way to report this. There is currently no mechanism that is widely published by governments or private institutions that is able to enforce swift action on such reports. Also, most of them consider these to be passive modes of misconduct. This is only a subset of the real challenges that the internet or web of tomorrow is going to intensify and further. If we visualize the same type of actions in the real society, we usually would end up with multiple complaints to the local police. For a better tomorrow, we need not only the law enforcers, technology creators, software developers but also informed netizens to replicate the real world society. This is a goal that can be achieved and still without too much of active administration.

By taking appropriate steps and making significant strides in them is what will lead to the sweeping, yet passive, changes in the current system. The most important aspects to achieve these changes are outlined below. Each should be a precursor for important research and development topic, by itself. To achieve them, requires the combined effort of software engineering, internet governance, net usage, censorship and cyber policing.

A. Bifurcation

This applies to any form of temporary or virtual information divergence. Though this would apply to mainly the entry points at the internet like search engines, social networks, directory or ecommerce listings. Although, this can be applied to larger sets of data, it can also be applied to logically related subsets. One example would be to deduce and apply context to each type of information searches. This would ensure that there is a clear demarcation of the type of data actually available. Another aspect would be to use behavior of the user to induce relevance to a user workflow. This aspect has to be implemented through configuration or centralized information exchanges between various participating or critical websites. In general, a repository that could hold every netizen's usage habits or information would be the per-fect sub-solution to bifurcation.

B. Prevention

There are enough mechanisms in place to loosely prevent out of track information to be added to a given set. This applies to simple form data, images, documents, videos, music and even multiple other data forms. But using more efficient robotic analysis, we should be able to prevent multiple situations such as terrorism, adult content, religious violence and racist remarks. These rules should be made available through an efficient regulatory body depending upon the type of the site. Also, this would help in reducing information chaos by not allowing unwanted, automated or non-context information to build up. Apart from this, this would clearly help in easing the process of Bifurcation of information. Regulated Prevention would also mean that we are able to provide a uniform barrier in the cyber society.

C. Reporting

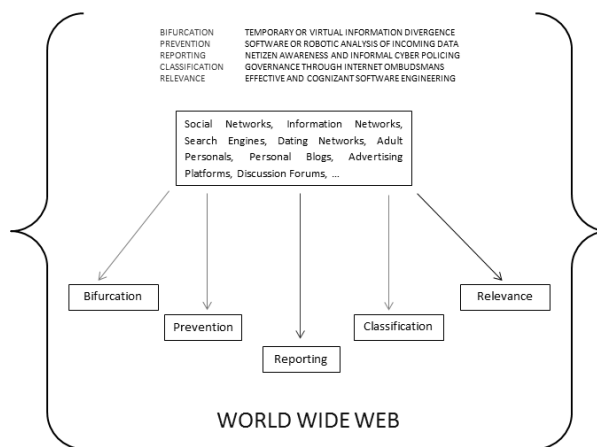
The simplest and most remedial measure to counter multiple information issues on the internet is to introduce more effective cyber policing. This needs to be enforced first by making available relevant information for reporting to the user. Also, an ordinary user must have ready tools and links to use reporting mechanisms. These mechanisms need to be appropriately placed at locations that are accessible and can be followed to create detailed descriptions for more analysis. Also, by creating a newer wing of current cyber reporting or analysis mechanisms that are swifter in action, whether online or on ground – we are removing fear from the users.

D. Classification

Governments all across the world, especially while forming cyber laws for the newer internet have to now follow stricter registration procedures for internet. This may include web startups, dotcoms, information sites, social networks or even academic websites. This will help in curbing non-genuine sites, fraudsters, scammers and malware sites. By allowing city, state, zonal, offline, online and national ombudsmen; we will be able to quickly enforce policing. Apart from these, by allowing a dedicated team of

internet information providers, that track live information; we will be able to classify information with much ease. Also, as time progresses, we should not allow a very easy domain registration and hosting process – that does not involve the ombudsman. The additional task of each ombudsman office would be certification, classification, ranking and maintaining historical data of its zone.

Information Engineering on Internet



E. Relevance

Software Engineers and Architects should use novel practices to create more relevance in the type of data that is accessible to the user. With stricter laws in place, we should provide only the right type of data operating within a context. Apart from this the engineers should be able to maintain a pool of intelligent information within their own applications or software. This will allow them to retire data sets that may not keep in line. They may also choose to provide this data to other developers. By automating most of the other tasks that are related to bifurcation, prevention, classification and reporting – they would be engineering more relevant software and thereby the internet. Software and Internet Architects will be required to co-ordinate and also obtain relevant certifications for their own properties. They also need to envision a way wherein the software implementers can implement all of the required policies with greatest ease. The software would also need to have standard integration points to quickly circulate information around common or related websites.

If we now quickly reflect on the original issues that we had with surfing the internet, we can classify the problems that are solved. These are as under the various aspects to curb information mess and also to curb anti-patterns from forming.

Bifurcation

Irrelevance, Non-Context Info, Malware, Adware, Ad Driven Linking...

Prevention

Underage, Adult, Dating, Religious Violence, Profanity, Racism...

Reporting

Bullying, Financial Frauds, Misleading Mails, Chain Mails, Terrorism, Frauds...

Classification

Pay for Click, Pay for Visit, Unsolicited Meeting, Phish-ing, Scam...

Relevance

Identity Theft, Misconduct, Irrelevance, Barriers, Racism, Threaten...

Now analyze the same situation and visualize with all these policies in place and running in the background; without your involvement – but with only greater user awareness. You are bound to have a safe, hassle free surfing experience and a more enjoyable cyber society. This one now duplicates the real one which we live in, but without taking away the positive freedom internet gives us.

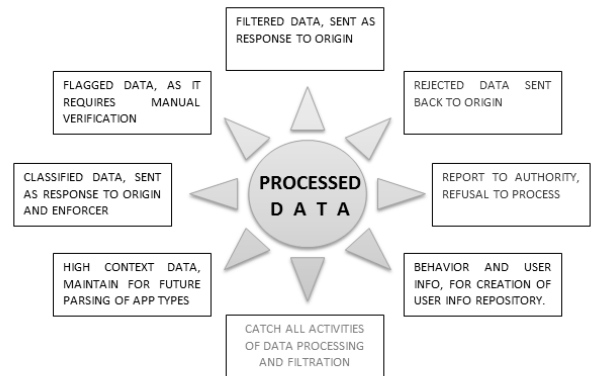
3 CASE STUDY

Let us apply this knowledge to a real world case study. This is out of genuine experience in usage of real world social networking sites. Through a social network or search engine, almost every user including under aged netizens have access to adult content including pornography, services and dating. Both of these tools or thought processes go hand-in-hand to achieve bifurcation, prevention, reporting, classification and relevance. The software titles mentioned below are only design descriptions or incomplete alpha versions for now.

A. CyberClean

The CyberClean component is an attempt to organize information via crawling and creating a mine of information. This should also contain divisions of data based on its relevance. CyberClean is a centrally available component that will crawl over the internet or provided configurable and secure data input streams which are then analyzed as per the standards. CyberClean is allowed to run at fixed intervals or continuously, depending on participation and volume of data. It will be designed to accept static or even live information. The incoming stream is decided on type or class of information that is to be filtered or processed.

The outcomes of data processing by CyberClean will be used to achieve various aspects of information sanity. The most important of them are classification, reporting and prevention. CyberClean will also set the data up for easier achievement of relevance and for rightful bifurcation.



< CYBER CLEAN – DISTRIBUTING FILTERED DATA TO VARIOUS PROCESSES >

• Prevention

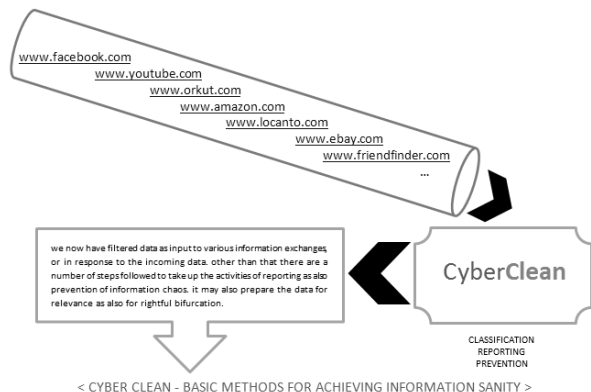
Through participation of data or web sites that are high traffic, CyberClean will have to deal with highly dynamic and static data as well. Processing of data from an array of sources such as social networks, adult networks, e-commerce sites, auction sites, advertising sites, and market-places will not allow non-context data to creep in. Also, it will allow rejection of such data using uniform laws across the sites. It will also remove the first barrier to achieve information sanity, which is influx.

• Classification

CyberClean's mainstay is to classify information from all such sources as mentioned above. CyberClean is able to create an intelligent repository by finding patterns of information, deriving context, finding deviations from specified rules for a given types and also enforcing laws. CyberClean is a tool for e-Governance for Internet as it will provide information directly to the internet ombudsman. This information includes health, types of data, level of chaos, user types, area ranges and also administrative inputs.

• Reporting

Reporting is also a very important outcome of CyberClean. CyberClean not only reports or detects various types of crime patterns, fraud patterns and any form of irregular information pattern. As mentioned earlier this would also be a form of self-initiated reporting by the tool. The information report would also specify any pas-sive form of violence such as sexual, abuse, religious and political. Reporting will also be an activity that could lead to an immediate adverse action without adequate warning of the creator. This is allowed as a result of accepted and enforced governance on the website or any sub-site of the main website.



B. Bunnybook

Bunnybook is an adult companionship and adult personals software that uses CyberClean's output of classification and prevention. Bunnybook solves an important problem of information chaos and influx building up on social networks, contact lists, ecommerce directory, auction sites and also on advertising sites. CyberClean provides the context information, labeled as 'adult companions', 'adult friends', 'adult personals', 'adult dating', 'escorts', 'adult services' and so on to the centralized distribution mechanism. This data, depending on governance or ombudsman specified laws can be shared or sold to Bunnybook. Bunnybook will be able to provide greater user experience, intelligence, relevant information, better business, better advertising and at the same create the right demand of information-supply of information chain. All this only when a relevant search is performed, else the data would have further added to a possibly less legal and less social a network.

• Bifurcation

Bunnybook is a concise example of rightful bifurcation of the internet. With greater usage and search on social networks or search engines, users will only contribute to further demarcation. Also, relevant bifurcation happens both ways in the case of data that is out of context. The out of context data now becomes more accessible for relevant searches on the internet. Bunnybook is also one example where a huge related mass of data becomes usable within a context, which otherwise would have made no sense or added to the chaos.

• Relevance

Bunnybook also reaffirms that with better software engineering skills we should be able to provide relevance. First, by using data that will be provided from Cyber-Clean. Next, by making sure that they are able to create chains and links to further such data. By linking related data, the software developer is making sure that we are creating a safer and a relevant data. Bunnybook will also make sure that the Software Engineer is able to make lay-ered security decisions on the information that is to be made accessible in the current session. In addition the software architect is able to make a

decision on obtaining relevant adult software related certifications. He is also able to make an easier decision regarding the required centralized policies to be implemented for this class of product. He is also able to make an easier decision regarding the required centralized policies to be implemented for this class of product.

• Prevention

Bunnybook by now becoming a relevant product on the internet that is known to provide quality adult friendship is preventing information chaos from its end. For all forms of dynamic or live data that was actually reported by CyberClean to not be acceptable on a social network, college network, friend network or even a dating site is now easily available or redirected to Bunnybook for relevant registrations. As this process is followed repeatedly, we are able to prevent data confusion by using other networks like Bunnybook. As this adult network grows wider and deeper with more features, it is also able to help in prevention by itself.

Using a combination of CyberClean, Bunnybook and using data on a popular social network, we are now able to create a more *Socially Acceptable and Aware Social Network*, and at the same time more relevant *Socially Acceptable Adult Software*.

4 CONCLUSION

The total number of internet users in the world is estimated at 39%, with only about 31% in the developing countries. With increase in education and basic computing facilities and knowledge, the developing countries will also start moving closer to the around 75% present usage in developed countries. This is while internet usage in developed countries moves closer to more than 95% of their population. While we ready ourselves for a total domination by internet and its associated technologies, we should also be ready to understand and embrace the far greater problems that it will continue to create. By recognizing these issues before hand, leveling out the current pitfalls and envisioning for a secure cyber society, the law enforcers should be able to make rapid strides and enforcement. Relevant and Aware Software Engineering can not only provide a basis for implementation of the thought processes, but by following a given direction they would be able to achieve Information Sanity faster.

The age of Internet policing and governance ushers in a new era but may also perplex the mind of an internet user. This bewilderment can only be minimized by creating awareness of laws and forming usage habits. It is also about time now that governments and citizens together agree that we need to create stricter cyber laws and enforce them. But by doing this, we need to still keep it a source of freewill, free thought and infinite knowledge.

5 ACKNOWLEDGMENT

The only acknowledgment I want to make here is of Wikipedia, the infinite source of relevant information.

6 REFERENCES

- [1] National Science Foundation, Beijing Jiaotong University, and the China Scholarship Council, "Internet Traffic Clasification" (references)
- [2] Anne Meredith Fulton, CyberSpace and Internet : Who will be the cyber police? (references).
- [3] Janne Lahtiranta, Kai Kimppa "Elderly People and Emerging Threats of the Internet and New Media"
- [4] Henry H. Perritt, Jr., "Is the Environmental Movement a Critical Internet Technology?"



Sumith Kumar Puri is an Independent Software Architect. He graduated as a Bachelor of Engineering [Information Science and Engineering] from Sri Revana Siddeshwara Institute of Technology, Bangalore, India.

He has nine years of progressive software development experience working with top software companies including Siebel, GXS, Huawei, Symantec, Misys and Yahoo. He is also a Sun Certified Java Programmer 1.4, 5.0; Sun Certified Business Component Developer 1.3, 5.0; Quest C Proficiency Certified, Quest C++ Proficiency Certified and Quest Data Structures Certified.

He is a Professional Member of the IEEE and IEEE Computer Society. Additionally, he is the member of the Association for Computing Machinery, India and the Computer Society of India. He has very deep interests and knowledge in Data Mining, Software Security and Advanced Programing in C/C++.