A REVIEW OF TECHNICAL ASPECTS OF INTERNET ANONYMITY, AND ITS TOOLS

[1]Parul, [2]Narender Sharma
[1]Research Scholar, [2]Assistant Professor
[1,2]Department of Computer Science & Engineering
[1,2]N.C. College of Engineering, Israna, Panipat, Haryana

Abstract: Communication is one of the most important medium through which people can share their beliefs, ideas, point of views and their knowledge. Privacy has become a major concern as innumerable devices all over the world communicate with each other on constant basis. Everyone wants to communicate securely over the network. Preserving privacy does not only mean hiding the content of messages, it also includes the protection of personal information of people communicating with each other. It's much like a physical postal letter, the simple technique of cryptography within a packet switched network that conceal the messages being sent, but it describes the identity of persons who is sending the letter and to whom the letter is being sent. Constant efforts are being made to secure privacy for e.g., Encryption is one such step which helps in hiding the identity of the persons communicating. Obviously it's a baby step in that direction and much more still needs to be done. To achieve the privacy is somehow related to anonymous communication over the network. This paper focuses mostly on acquaintance of anonymity systems, technical aspects of Internet Anonymity and some sort of the common tools available on the Internet to fulfill the user requirements, and it will also describe the most popular tools in use.

Keywords: Anonymity, Pseudonymity, Anonymous Communication, Internet, TOR, Tools

1. Introduction:

Privacy is a most desirable feature of modern society. There is no exact definition for privacy, although there are a few alternatives who suggested the term privacy with its potential meaning. The basic *definition of privacy* is assuring that sensitive information will be remain secret and access limited to the appropriate persons. An alternate definition from the Merriam Webster is "Privacy is the quality or state of being apart from company or observation or freedom from unauthorized intrusion." Modern infrastructure of communications includes Internet as basic entity that enables us to connect people around the world at just finger tip. For further connection of devices a unique address is required over these networks of Internet i.e., Internet address. Whenever one communicate using the Internet, one leaves "tracks" which identify where the message (packet) came from and was going to. IP addresses are allocated to individuals or companies by their Internet Service Providers (ISP), who keep this information; thus ensure that it is possible to track every Internet communication down to an individual or at the very least, organization. Data packets travelling over the Internet has its source and receiver address. If that packets travel with the largest structure of networks i.e., routers, switches, gateways etc, it reveals some services of network instance who is talking to whom though IP addresses. In some cases, privacy is really is not required for communications but it is required in the form of necessity. Privacy of communications can be easy to achieve by using wide variety of tools and well-known encryption techniques.

2. Contributions

Through this project we strived to introduce the idea of anonymity systems and did the following to make our contribution. First, we looked at the concept of Anonymity and Pseudonymity Systems and then explained the basics of both Systems. We discussed about Technical Aspects of Internet Anonymity and techniques associated with it. Then we presented verified diverse tools of anonymity systems with their working explanation. Also, we presented methods of anonymity analysis for a variety of data sets i.e., data, images, registration plates, in Medical fields.

3. Anonymity and Pseudonymity

These are the kind of techniques for ensuring Privacy. Before going into more details of Anonymity system, Lets have the meaning of both Anonymity and Pseudonymity.

Anonymity is the condition of being anonymous over the network whereas Pseudonymity refers to false identity of someone over the network. According to Cambridge dictionary, Anonymity refers to the situation in which someone's name is not given or known. Anonymous communication is handy in several kind of situations or even able to work in unfriendly environments, e.g., wrapping of identity of an analyst who is examining for a malware to protect from counter attack, permitting the expert of security to circumvent of network security policies during penetration testing of the network and services, it consent to commuters to their home network services through the Internet even network is not connected to VPN, it preserve data security in particular networks. Although different systems, models and protocols of anonymity have been developed, this is responsible to provide unlinkability or untraceability between the received messages and their true senders and between the sent messages and their true receivers. Most pseudonym holders use pseudonyms system to use false identity, because they wish to remain anonymous, but anonymity is not an easy task to achieve, and is often having legal issues. Real anonymity system doesn’t need ability of link in any way, opposite of it result such as an attacker's inspection of the pseudonym holder's message provides no new information about the holder's exact and true name that will also hide his/her identity.

Most Websites that have services to offer pseudonymity keep track of information about their users. These sites are often at risk to unauthorized intrusions into their non-public database systems i.e., security breach into private database systems. According to a study of a Web dating service and a pseudonymous remailer, performed by the University of Cambridge researchers discovered that the systems used by these Web sites to protect user data could be easily compromised, even if the pseudonymous channel is protected by well built-in encryption. In general, the protected pseudonymous channel exists inside a broader framework in which multiple vulnerabilities exist. Pseudonym users ought to have clear image in mind that, given this state of Web security engineering, their true names can be exposed at any time. [5]

Today, more than half of Internet users are concerned for the privacy of their personal data on the Internet. As shown in below chart based on Survey conducted by Attentiv states different opinions of people about anonymity and according to that 59% (Internet users) people believe that Internet anonymity can’t be achieved and only 37% people (Internet users) believes that it’s possible to use the Internet anonymously.[11]

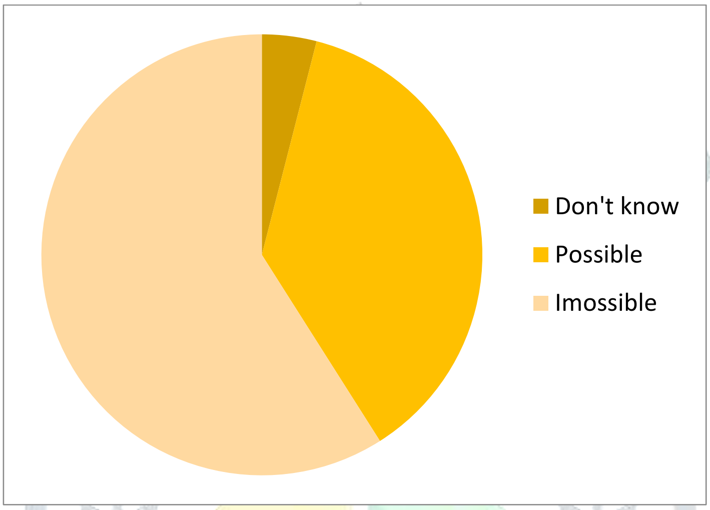


Figure: Result of online survey on “Is it possible to be anonymous online? from group of Internet users”

4. Technical Aspects of Internet Anonymity

Many Americans presume that the Internet mechanically provides an exact level of anonymity. (In different countries, this can be not so; Chinese people, for instance, expect their messages to be watched and censored) most vital messages sent over the Internet square measure encrypted, therefore the content of the message can't be simply deciphered. However, if proper measures to protect the information are not placed, each action a person takes on the Internet may be copied back to your IP address. A "man in the middle" will simply intercept your IP address (and thus your approximate physical location) and may, if he is well equipped can penetrate your laptop and can get access to private & confidential information that is not the sole method of distinctive Internet users.

There are three main ways of obtaining anonymity on the Internet:

- Proxy services are the easiest among all to use, but that are less secure.
- Anonymous networks and anonymizing software, such as I2P and the Tor Project, are gaining popularity.
- Peer-to-peer networks like Freenet are much protected, and are also slowly gaining membership.

Even with the foremost advanced anonymizing software system, there are still ways in which your data can get leaked to the web. Nothing is confidential or personal be it the information or personal details if it is transferred or conversed via Internet by collecting the information from Internet, one can get complete personal data and can even trace the location of the person. It’s almost impossible to hide online communication as based on information we share online, one can track any info required by that person. A talented huntsman can use these details to make a complete profile.

Another drawback of this is prolonged existence and index-ability of on-line data. Whereas physical data (information sheets, deeds, checks, letters, etc.) tend to exist solely as one copy archived somewhere (possibly remote) and are capable of being lost easily, on-line data stays around. It is totally difficult to utterly wipe off any trace of the existence of data on-line, which can be saved on multiple backup servers or mirrors. And with indexing bots (such as Google's "spiders") sifting through the Internet daily, it is more and more possible that an easy search will turn up this information. Somebody with enough determination and smart searching tools will notice information you thought long lost.

4.1 Browser Fingerprinting

Browser Fingerprinting is a technique which is used to track users on websites just from their browser configuration and settings information they make visible to websites. It is helpful in successfully tracking users without cookies or using any other tracking standards. This is basically similar to identify a person by visual inspection. It's fine if you visit a website once or use internet to converse for a single time but if you keep revisiting the same site on multiple occasions, fingerprinting can recognize the person even if no information has been revealed.

Panopticlick, it is a research project of the Electronic Frontier Foundation (EFF). It is a part of an effort to demonstrate the problem with tracking techniques, and help get stronger privacy protections for everyone. It is a research project designed to better expose the tools and techniques of online trackers and analysis the worth of privacy add-ons.

When you visit an Internet site, you are permitting that web site to access a lot of information regarding your computer's configuration. Combined, this info will produce a form of fingerprint, a signature that would be used to establish you and your computer. Some corporations use this technology to do to spot individual computers.

In 2010, EFF launched Panopticlick, a research project to investigate how unique each browser is. They gathered information about the configuration and version information from your operating system, your browser, and your plug-ins, and compared it to their database of many other Internet users' configurations. Then, they generated a uniqueness score — letting you see how easily identifiable you might be as you surf the web.

In 2015, they upgraded Panopticlick with a new feature: tracker blocker testing. More than half of Internet users are using privacy add-ons and other tools to block trackers. There are thousands of free add-ons, created by developers all over the world that you can install to personalize your browser. It includes tools like AdBlock, Ghostery and Disconnect.

But how well do these add-ons truly protect users from persistent tracking? Their new version of Panopticlick researches both. They analyze how well you are protected against online tracking by checking the privacy protections you have in place. The test simulates loading of a visible ad that performs tracking, an invisible script that performs tracking, and a site that looks superficially like a tracker but actually has committed to honor Do Not Track.

Even if your privacy add-ons are working well, your browser will still be vulnerable if your browser fingerprint is unique. So they also analyze the uniqueness of your browser and let you know how it stacks up to other visitors they have experiment recently. They will generate a report about your tracker protections and browser fingerprint for your own use, and they will include anonymous results from your test in their larger research report. Running tests on Panopticlick both gives you this information about your own browser, and also helps EFF use statistical methods to evaluate the capabilities of Internet tracking and advertising companies, and the best forms of protection against tracking without consent.

4.2 Proxy Services

Proxy services work by sending requests from one user through a different computer. In this way, a user (or the endpoint website) is seen only to be communicating with the proxy server, and (as long as the message is encrypted) anyone monitoring the message from either end can't determine what the other endpoint is. This is the simplest and easiest way of securing anonymity on the Internet. To an outside observer, the message gives nothing away.

This is incredibly useful on both ends. First, someone trying to track down your IP address from a message you sent over the Internet, will see that message as having come from the proxy server not from you. The proxy server can be anywhere in the world and betrays no information about who you are, where you are, or what your IP address might me.

On the other end, if someone (say, a totalitarian government) is monitoring your Internet usage, all they will see of your activity is communication with the proxy server. If this is not a blocked address, the government has no reason to block your communication, and so you can access banned content rather easily.

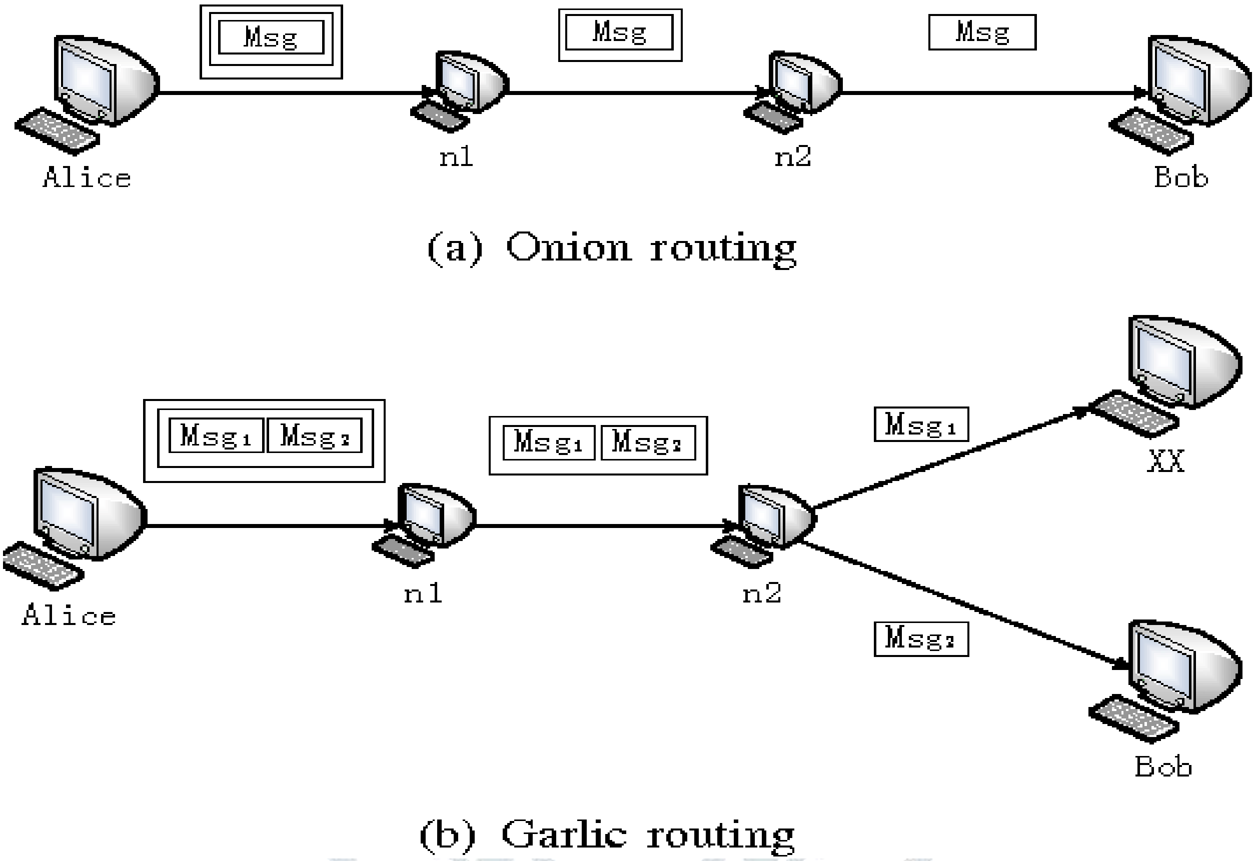
Of course, once the address is determined to be a proxy website, it will also be banned, but there are always more and changing IP addresses of different proxy sites, allowing for the free (if illegal) transfer of information worldwide through the Internet. Proxies are far less secure than other anonymizing methods, however, because they have single-point weakness. If someone manages somehow to gain access to the proxy server, the server will give up all the important information it has been concealing. As mentioned above, it can also be easily spot-banned, and with a single censorship, an entire proxy route becomes unusable. Distributed networks like I2P or Tor work around this to make a much safer proxy-chain service.

4.3 TOR and Deep Web

I2P (Invisible Internet Project) and projects like it are still being developed and are not in wide use. However, I2P makes use of encryption that makes it possible to hide a user's IP address from the recipient of a message, or a third-party. Tor is a similar and more widely known project. The Invisible Internet Project (I2P) uses garlic routing ("The invisible Internet project") which is an extension of onion routing. [8] With its ability to provide effective routing anonymity, Tor allows users an unprecedented opportunity to be truly anonymous. Traditional websites host a large body of relatively static content. Search engines indexing the web find the site through links to it from other sites, then parse the content and make it available for locating via web searches. While the data traveling across secured HTTPS links is impossible to read, it's easy to eavesdrop the general information about the packets, like their sender address and receiver address. With this information, it's possible to locate the physical systems on either side of the exchange.

TOR removes this ability. Tor, or The Onion Router, uses a large number of nodes, routing packets across the globe between random nodes until choosing to leave at an exit node and finally be delivered by the exit node to the final destination. While it's possible to see a connection from a client machine to a Tor entrance node, it's effectively impossible to follow the packet as it traverses the Tor network. This removes the possibility of establishing a link between the client and the server they are requesting. Every user of Tor can choose to run the software in either client mode or server mode. If they run in server mode, their computer can be used as one of the nodes along the routing chain. This means the nodes can be spread all over the world and knowing the location of one node tells you nothing about the rest of the chain.

With the Tor network between clients and servers, the possibilities of anonymity are not limited to the clients. A number of websites and servers have sprung up based around the opportunities of anonymity. These servers are equal to a real black market of the Internet. There you can find almost anything you want, if you have the desire and the right contacts. This is part of the deep web, as it's not trolled by the many web spiders of the various search companies, and not listed on any of their associated indexes. There is no way to browse websites, except via individual links, commonly acquired by word of mouth between online contacts, though some listings of major sites can be found on existing shallow-web sites.



(a) Onion routing

(b) Garlic routing

Figure Routing Methods for Anonymous communication

5. Data Anonymization Tools

As we already discussed about Anonymity System. We are now moving to next crucial topic of this paper i.e., tools for Data Anonymization.

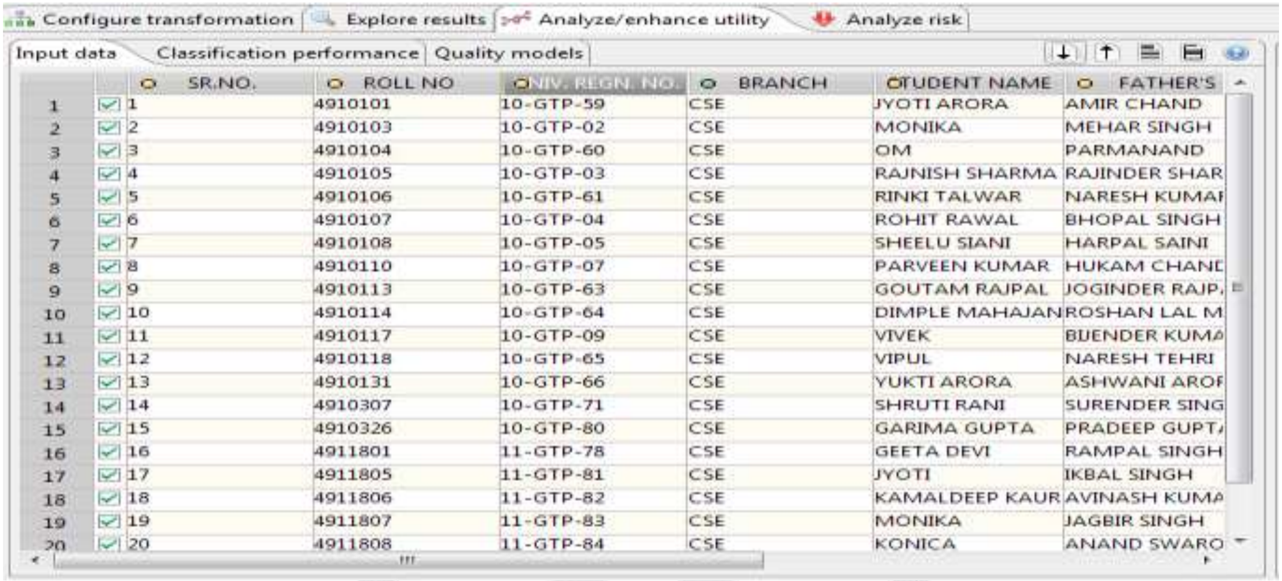
Data Anonymization is a process which intends to protect our privacy. The main idea behind Data Anonymization is to encode or remove personally identifiable information from datasets. In this manner, people from whom the data was collected remain completely Anonymous. Data Anonymization is mostly used in the medical field and in some cases to perform data-sensitive statistics analysis. In case you want to disclose the information you are performing on anonymous, below listed tools can be used:

Data Anonymization Software

5.1 ARX Data Anonymization Tool

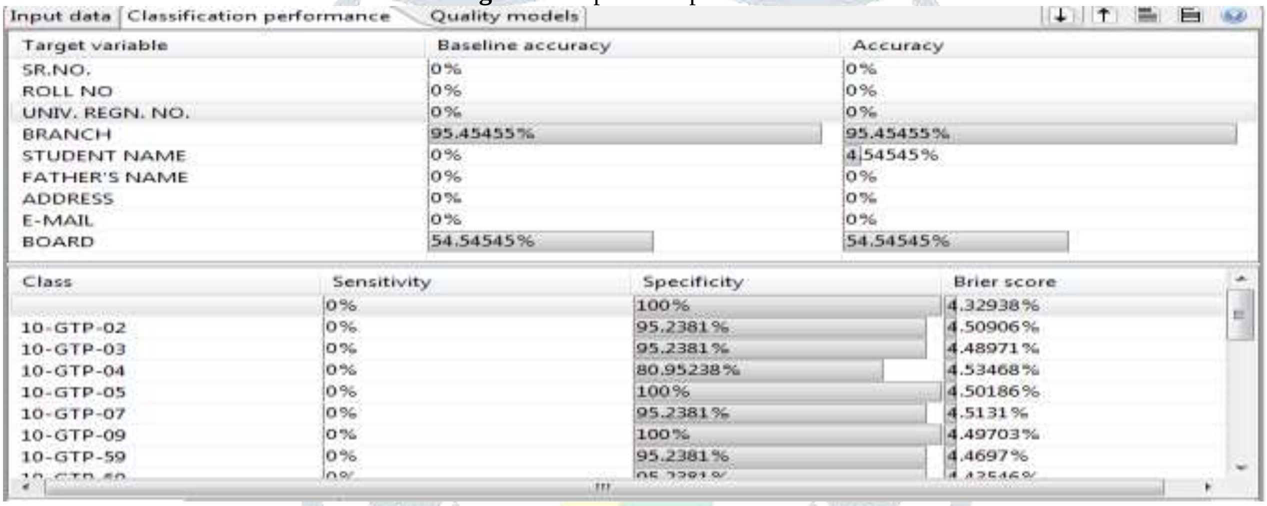
ARX is an openly available software package for the anonymizing sensitive personal data. The tool transforms data sets into syntactic privacy model that moderate attacks leading to privacy breaches. This tool removes direct identifiers similar to names from data sets and adds further constraints on indirect identifiers like email addresses or phone numbers. The tool furthermore provides integral data import facilities for relative databases (MS SQL, DB2, SQLite, My SQL), MS-Excel and CSV files.

ARX is a best option of a cross-platform graphical tool, which maintains data import & cleansing, wizards for creating transformation rules, spontaneous ways that during which for craft the anonymized dataset to your desires and visualizations of data utility and re-identification risks. You can get ARX from their authorized website, as mentioned in reference section. [6]



		SR.NO.	ROLL NO	UNIV. REGN. NO.	BRANCH	STUDENT NAME	FATHER'S
1	☑	1	4910101	10-GTP-59	CSE	JYOTI ARORA	AMIR CHAND
2	☑	2	4910103	10-GTP-02	CSE	MONIKA	MEHAR SINGH
3	☑	3	4910104	10-GTP-60	CSE	OM	PARMANAND
4	☑	4	4910105	10-GTP-03	CSE	RAJNISH SHARMA	RAJINDER SHAR
5	☑	5	4910106	10-GTP-61	CSE	RINKI TALWAR	NARESH KUMAI
6	☑	6	4910107	10-GTP-04	CSE	ROHIT RAWAL	BHOPAL SINGH
7	☑	7	4910108	10-GTP-05	CSE	SHEELU SIANI	HARPAL SAINI
8	☑	8	4910110	10-GTP-07	CSE	PARVEEN KUMAR	HUKAM CHANE
9	☑	9	4910113	10-GTP-63	CSE	GOUTAM RAJPAL	JOGINDER RAJP.
10	☑	10	4910114	10-GTP-64	CSE	DIMPLE MAHAJAN	ROSHAN LAL M
11	☑	11	4910117	10-GTP-09	CSE	VIVEK	BIJENDER KUMA
12	☑	12	4910118	10-GTP-65	CSE	VIPUL	NARESH TEHRI
13	☑	13	4910131	10-GTP-66	CSE	YUKTI ARORA	ASHWANI AROF
14	☑	14	4910307	10-GTP-71	CSE	SHRUTI RANI	SURENDER SING
15	☑	15	4910326	10-GTP-80	CSE	GARIMA GUPTA	PRADEEP GUPT/
16	☑	16	4911801	11-GTP-78	CSE	GEETA DEVI	RAMPAL SINGH
17	☑	17	4911805	11-GTP-81	CSE	JYOTI	IKBAL SINGH
18	☑	18	4911806	11-GTP-82	CSE	KAMALDEEP KAUR	AVINASH KUMA
19	☑	19	4911807	11-GTP-83	CSE	MONIKA	JAGBIR SINGH
20	☑	20	4911808	11-GTP-84	CSE	KONICA	ANAND SWARO

Figure: Sample of imported data



Target variable	Baseline accuracy	Accuracy
SR.NO.	0%	0%
ROLL NO	0%	0%
UNIV. REGN. NO.	0%	0%
BRANCH	95.45455%	95.45455%
STUDENT NAME	0%	4.54545%
FATHER'S NAME	0%	0%
ADDRESS	0%	0%
E-MAIL	0%	0%
BOARD	54.54545%	54.54545%

Class	Sensitivity	Specificity	Brier score
	0%	100%	4.32938%
10-GTP-02	0%	95.2381%	4.50906%
10-GTP-03	0%	95.2381%	4.48971%
10-GTP-04	0%	80.95238%	4.53468%
10-GTP-05	0%	100%	4.50186%
10-GTP-07	0%	95.2381%	4.5131%
10-GTP-09	0%	100%	4.49703%
10-GTP-59	0%	95.2381%	4.4697%
10-GTP-60	0%	95.2381%	4.43546%

Figure: Resultant image of Data Anonymization

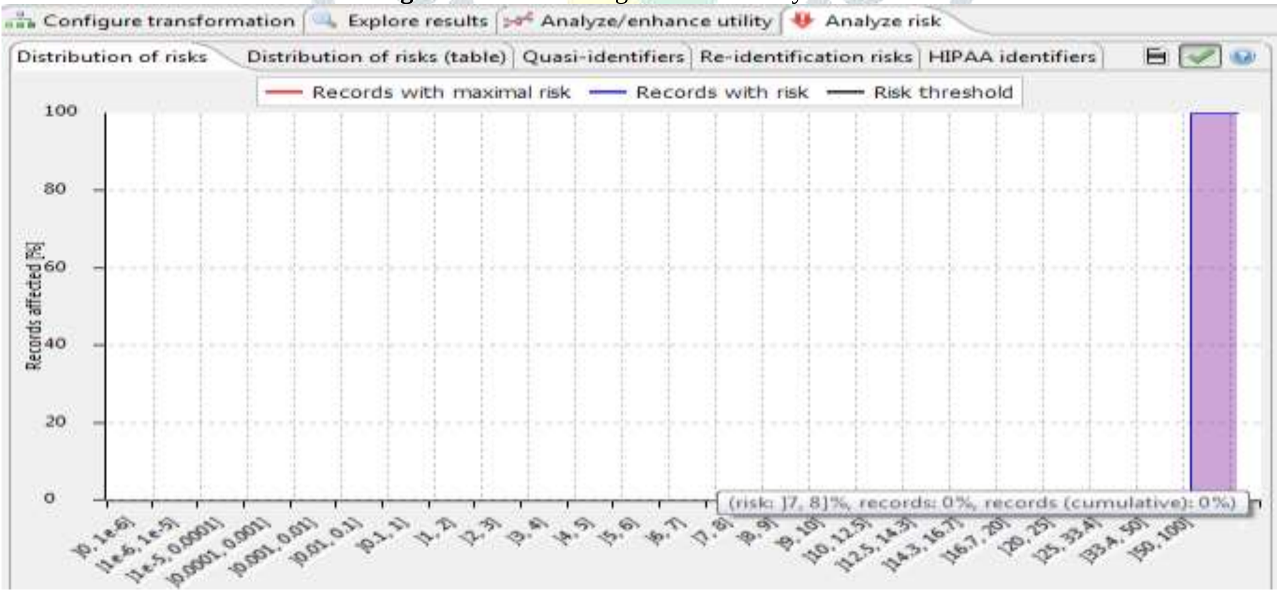
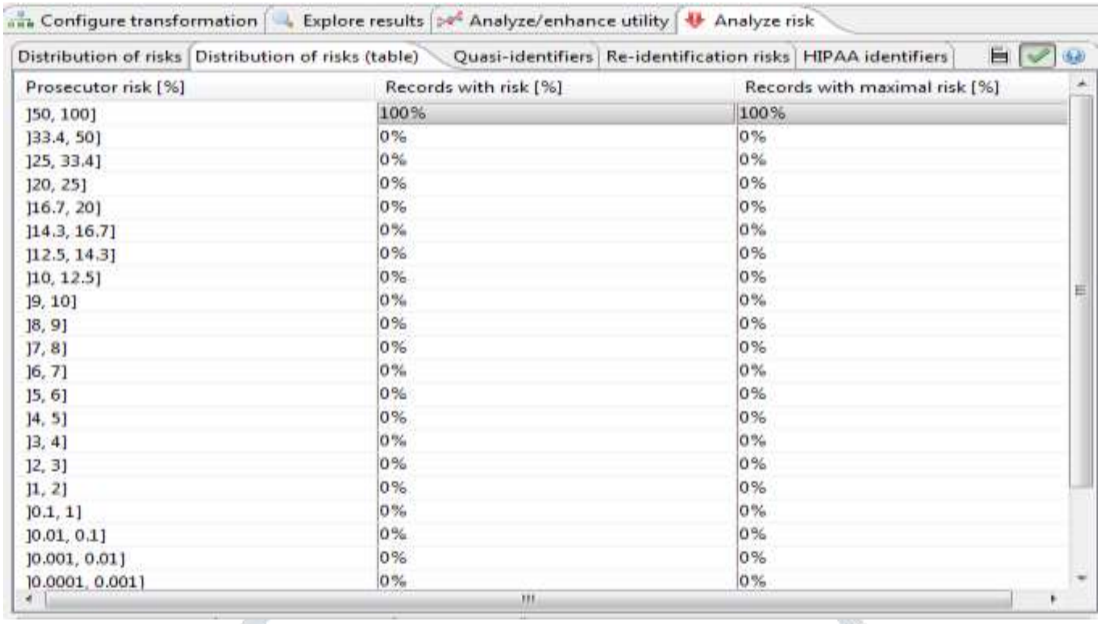
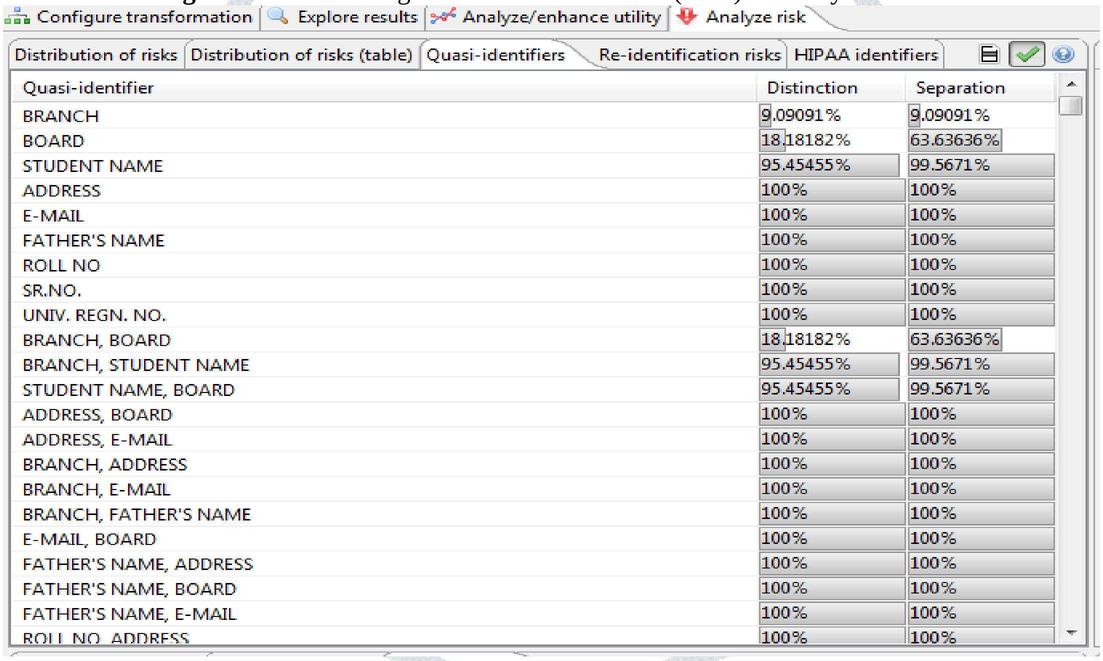


Figure: Resultant image of distribution of risks.



Prosecutor risk [%]	Records with risk [%]	Records with maximal risk [%]
]50, 100]	100%	100%
]33.4, 50]	0%	0%
]25, 33.4]	0%	0%
]20, 25]	0%	0%
]16.7, 20]	0%	0%
]14.3, 16.7]	0%	0%
]12.5, 14.3]	0%	0%
]10, 12.5]	0%	0%
]9, 10]	0%	0%
]8, 9]	0%	0%
]7, 8]	0%	0%
]6, 7]	0%	0%
]5, 6]	0%	0%
]4, 5]	0%	0%
]3, 4]	0%	0%
]2, 3]	0%	0%
]1, 2]	0%	0%
]0.1, 1]	0%	0%
]0.01, 0.1]	0%	0%
]0.001, 0.01]	0%	0%
]0.0001, 0.001]	0%	0%

Figure: Resultant image of distribution of risks (table) of Analyze Risks.



Quasi-identifier	Distinction	Separation
BRANCH	9.09091%	9.09091%
BOARD	18.18182%	63.63636%
STUDENT NAME	95.45455%	99.5671%
ADDRESS	100%	100%
E-MAIL	100%	100%
FATHER'S NAME	100%	100%
ROLL NO	100%	100%
SR.NO.	100%	100%
UNIV. REGN. NO.	100%	100%
BRANCH, BOARD	18.18182%	63.63636%
BRANCH, STUDENT NAME	95.45455%	99.5671%
STUDENT NAME, BOARD	95.45455%	99.5671%
ADDRESS, BOARD	100%	100%
ADDRESS, E-MAIL	100%	100%
BRANCH, ADDRESS	100%	100%
BRANCH, E-MAIL	100%	100%
BRANCH, FATHER'S NAME	100%	100%
E-MAIL, BOARD	100%	100%
FATHER'S NAME, ADDRESS	100%	100%
FATHER'S NAME, BOARD	100%	100%
FATHER'S NAME, E-MAIL	100%	100%
ROLL NO, ADDRESS	100%	100%

Figure: Resultant image of data Quasi-identifiers of Analyze Risks.

5.2 Anonymizer/ Image Data Anonymization

Anonymizer software anonymizes pictures using cutting-edge detection and blurring technology developed by Eyedea Recognition. This Software detects faces and automobile registration plates at a large scale by applying blurring filters which makes the faces unrecognizable and the automobile registration plates undecipherable.

Anonymizer software development kit is an independent C/C++ library for Windows and LINUX platforms. The core functions of library are to anonymize image files, JPEG buffers and RGB image buffers. Anonymizer is meant to be used for anonymizing of street-view photos, webcams photos, user submitted photos or other photos where you want to safeguard privacy. [7]

Basically this software focuses on two important mechanisms:

a. **Face detection:** This algorithm analyzes the faces in target images and makes those images in the form of Anonymization. This software is so powerful that it can even detect faces on images with low resolution.

b. **License Plate detection:** This algorithm is so accurate that it can detect all 1-line and 2-line EU-type registration plate.

Sample demo of this software is as:

Figure: Sample image for face detection

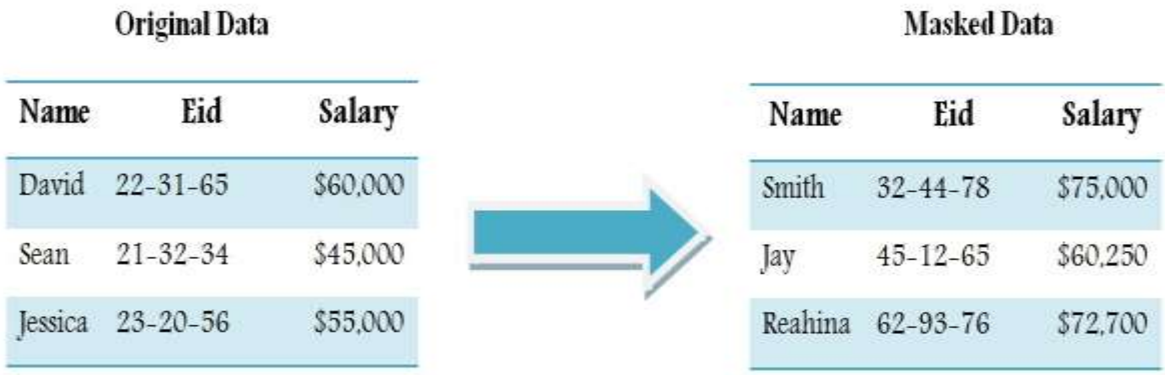
Figure: Resultant image of face detection Software.

Note: Web demo for image Anonymization is available on its official website and trial software (demo) version is also available for different platforms which are limited for 30 images. After using demo software, you are required to have a license key to use the product.

5.3 **Camouflage's CX-Mask**

CX-Mask removes the sensitive data hindering test, outsourcing, and analytics. The tool de-identifies sensitive data, and retains the realism and practicality of the initial information set. The information classes these tools will work with include names, addresses, credit cards, SSN/SIN, phone, and more.

Data masking, also cited to as data de-identification, pseudonymization, anonymization or obfuscation, could be a method of protecting sensitive data substitution original data with fictitious however realistic data. By masking information, organizations change information to be safely used in things wherever real information is not required.



Original Data

Name	Eid	Salary
David	22-31-65	$60,000
Sean	21-32-34	$45,000
Jessica	23-20-56	$55,000

Masked Data

Name	Eid	Salary
Smith	32-44-78	$75,000
Jay	45-12-65	$60,250
Reahina	62-93-76	$72,700

Figure: Data masking replaces original information with fictitious, realistic information.

Common samples of business processes that need access to realistic information, however not the particular data include:

Application development and testing – Application developers and QA groups want information to create positive applications work after they roll into production environments.

Outsourcing – Firms typically accept outside service suppliers and suppliers that require access to information for analysis, analysis, training, testing or development.

Coaching – Coaching departments need information to populate coaching systems that customers or partner's access.

Business intelligence (BI) and analytics – business analysts and researchers have to be compelled to combination and analyze information.

Organizations mask information for 2 primary reasons:

5.3.1 Protect sensitive data: It is common observe for organizations to repeat data from production systems to be employed in alternative non-production environments. These copies of sensitive data increase the potential attack surface and unnecessarily expose sensitive data to workers who might not be approved to access that data. Data masking removes sensitive data from non-production environments, reducing the prospect of data breach. It is a significant tool in an extremely multi-layered data security strategy.

5.3.2 Compliance: Several data privacy and protection rules specialize in safeguarding personal data like in Personally Identifiable Information (PII), health records and monetary data. Several of those rules require limiting access to sensitive information supported a need-to-know basis. For instance, the EU General Data Protection Regulation (GDPR) introduces data reduction and pseudonymization as key data protection principles organizations must follow. Alternative regulations, like PCI DSS Requirement 6.4.3, specifically prohibit the utilization of production data for check and development. Data masking is a vital tool to assist organizations avoids unwanted information access, scale back sensitive information exposure, and improve their compliance posture.

The purpose data masking is to supply data that appears and acts just like the original data, however lacks the sensitivity of the initial information. This manner the mask data does not create a risk of exposure or unauthorized access. It conjointly permits organizations to use fewer security controls for the covert information repositories and to cut back the scope of compliance audits.

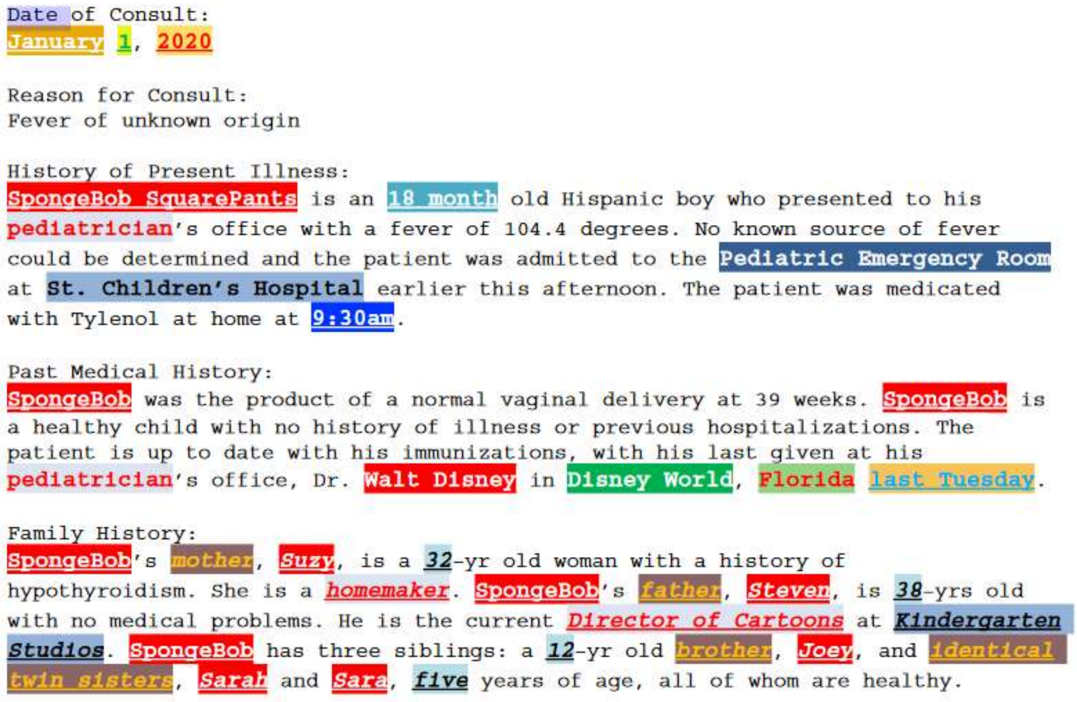
5.4 NLM-Scrubber

NLM-Scrubber is a new, free medical text de-identification tool. The software package is currently in its early beta stage; however you will be able to already attempt it out if you're curious. The good news is that its developers will unleash a non-beta version of the tool. NLM-Scrubber is to be chiefly used for de-identifying medical documents.

Here's the way to run NLM-Scrubber:

1. Download NLM-Scrubber from their official Web site. [12]
2. Produce a configuration file employing a text editor
3. Open a command-line interface > adjust the present directory to wherever you downloaded or stimulated NLM-Scrubber
4. Use the subsequent command to run NLM-Scrubber: scrubber.exe.

A design of imaginary medical Reports with Color Coded observations Using Visual Tagging Tool (VTT) is as following:



Date of Consult:
January 1, 2020

Reason for Consult:
Fever of unknown origin

History of Present Illness:
SpongeBob SquarePants is an 18 month old Hispanic boy who presented to his pediatrician's office with a fever of 104.4 degrees. No known source of fever could be determined and the patient was admitted to the Pediatric Emergency Room at St. Children's Hospital earlier this afternoon. The patient was medicated with Tylenol at home at 9:30am.

Past Medical History:
SpongeBob was the product of a normal vaginal delivery at 39 weeks. SpongeBob is a healthy child with no history of illness or previous hospitalizations. The patient is up to date with his immunizations, with his last given at his pediatrician's office, Dr. Walt Disney in Disney World, Florida last Tuesday.

Family History:
SpongeBob's mother, Suzy, is a 32-yr old woman with a history of hypothyroidism. She is a homemaker. SpongeBob's father, Steven, is 38-yrs old with no medical problems. He is the current Director of Cartoons at Kindergarten Studios. SpongeBob has three siblings: a 12-yr old brother, Joey, and identical twin sisters, Sarah and Sara, five years of age, all of whom are healthy.

Figure: Resultant image of NLM-Scrubber with Color Coded observations Using Visual Tagging Tool

6. Acknowledgement

We take this opportunity to thank Er Jagtar Singh, Head of Department of Computer Engineering for making essential facilities available for us.

References:

[1] “Anonymity Meaning in the Cambridge English Dictionary.” *Gender Pay Gap Definition in the Cambridge English Dictionary*, dictionary.cambridge.org/dictionary/english/anonymity.

[2] Kamat, P., Zhang, Y., Trappe, W., Ozturk, C. Enhancing source-location privacy in sensor network routing. In: Proc. 25th IEEE International Conf. on Distributed Computing Systems, 2005.

[3] Du Pont, George F. (2001) The Criminalization of True Anonymity in Cyberspace Archived 2006-02-21 at the Wayback Machine. 7 Mich. Telecomm. Tech. L. Rev.

[4] Post, David G. (1996). Pooling Intellectual Capital: Thoughts on Anonymity, Pseudoanonymity, and Limited Liability in Cyberspace Archived 2007-09-27 at the Wayback Machine.. *University of Chicago Legal Forum*

[5] Clayton, R.; Danezis, G.; Kuhn, M. (2001). "Real World Patterns of Failure in Anonymity Systems" (PDF). Lecture Notes in Computer Science. Lecture Notes in Computer Science. 2137: 230–244. doi:10.1007/3-540-45496-9_17. ISBN 978-3-540-42733-9.

[6] “ARX, Data Anonymization Tool.” *ARX – Data Anonymization Tool*, arx.deidentifier.org/.

[7] Řihák, Jan, et al. *Image Data Anonymization - Eyedea Recognition s. r. o.*, www.eyedea.cz/image-data-anonymization/.

[8] Winkler, Stephanie, and Sherali Zeadally. "An Analysis of Tools for Online Anonymity.

[9] “Panopticlick, a Research Project of EFF.” *Panopticlick | Self-Defense*, panopticlick.eff.org/.

[10] “The Best 4 Data Anonymization Software to Use.” *Windows Report - Windows 10 and Microsoft News, How-to Tips*, Feb. 2017, windowsreport.com/data-anonymization-software/.

[11] “The Anonymity Impossibility: Stats, Surveys, and Figures.” *Attentiv*, 30 Dec. 2015, attentiv.com/anonymity-impossibility/.

[12] “NLM-Scrubber Download Page.” *U.S. National Library of Medicine*, National Institutes of Health, scrubber.nlm.nih.gov/files/.

[13] “NLM-Scrubber Annotation Page.” *U.S. National Library of Medicine*, National Institutes of Health, scrubber.nlm.nih.gov/annotation/.

[14] “Anonymity on the Internet.” *Anonymity on the Internet*, people.dsv.su.se/ ~jpalme/society/anonymity.html.

[15] “Anonymity on the Internet Must Be Protected.” *Warren and Brandeis, "The Right to Privacy"*, groups.csail.mit.edu/mac/classes/6.805/student-papers/fall95-papers/rigby-anonymity.html.

[16] “Onion Routing.” *Wikipedia*, Wikimedia Foundation, en.wikipedia.org/wiki/Onion_ routing.

[17] Dhankani, Manish, et al. "Anonymous Communication System based on Onion Routing." *International Journal of Computer Applications* 121.23 (2015).