ISSN No. 0976-5697

Volume 3, No. 2, March-April 2012

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Framework for Profiling User Instant Messaging Behaviors on Anonymous Peer to Peer Networks

Longe, O.B.
Fulbright Fellow & Scholar-In-Residence
International Centre for Information Technology & Development
Southern University System, Baton Rouge, Louisiana, USA
longeolumide@fulbrightmail.org

Abstract- The sheer size of e-mail and instant messages received on daily basis by clients using mobile applications has created a usage scenario where users have to be conscious of their safety and manage information effectively in order to derive benefits from online interaction. This is moreso in organizations where e-mails and instant messages are frequently used as a means of communication through anonymous peer-to-peer systems (P2P). Using Use-Case schemes, we present a framework for profiling user behaviour on anonymous peer to peer networks. The framework presented is suggestive of employing data mining techniques to analyze entire sets of active and offline e-mails and instant messages sent and received by individual users. The implementation of this framework is expected to yield a recommender system for prioritizing e-mails and instant messages based on usage patterns between contacts and groups to which users belong.

Keyword*:* Behaviour, chat, Data mining, e-mail, instant messages and recommender systems.

I. INTRODUCTION

One of the reasons why anonymity is a desirable trait in peer-to-peer networks over standard P2P systems is because of its ability to prevent DOS attacks against particular hosts. For instance, in standard P2P networks it is possible for attackers to map IPs and bring down any nodes through dedicated attacks tactics. In an anonymous P2P network, mapping nodes to host is practically impossible thus preventing attacks through conventional means [1].

Whistle-blowing within an organization is another scenario where anonymity can be an advantage. In this case, it will be difficult for organizations to identify a whistle-blower within the network and thus not able to prosecute such actions. Finally, an anonymous P2P construct can allow applications in which anonymity is desired to be created easily; one such example is electronic voting. In such an application anonymity is needed to prevent the vote being tied to the voter.[2]

However, with increase in unethical conducts on anonymous P2P systems which translates to similar offline behaviour, it has become imperative to devise means through which users can be identified on these networks.

Since P2P networks offer some forms of anonymity that can aid unethical conducts on the internet [3], behavioural profiling has been muted as one effort that can assist in stemming the challenges of unethical use of online facilities, research is warranted into developing system that assist in drawing inferences from patterns of on online conduct [4].

This will aid behavioural scientists in their quest to understand human conduct from diverse socio-ethnic background and be able to predict possible offline behaviours from available information. Furthermore, since users also go online for economic, social, educational and entertainment and other related activities, profiling will assist industry researchers to predict products and services that can be targeted to particular age groups.

For instance, if from internet usage profile, it is discovered that a user has the habit of sending dubious mails or messages, it is likely that the person is naturally criminally inclined. It can also mean that the person is not gainfuly employed or move around with other criminals.

A. Standard and Anomy nous Peer To Peer Systems:

Peer to peer network offers usage platforms in which the users and their nodes are pseudonymous by default. On P2P networks, it is possible to share files, stream media (audio and video) and create discussion forums. The primary difference between standard P2P and anonymous P2P networks is in the routing method of their respective network architecture. Anonymous P2P networks enable unlimited flow of different types of information, both legal and illegal [5]. Interestingly, the P2P community's interest in anonymous systems has increased rapidly in recent years. Some of the reasons for this increase include the difficulty in obtaining approval for publication under circumstances of official censorship of government and the mixed bag of sharing copyrighted files illegally and having unfettered access to spurious web contents such as internet pornography and illegal websites [6].

An attribute of a P2P network is that all clients provide computing resources that include bandwidth, processing capability and storage space. Thus, the system is scalable to system demands occasioned by node expansion. On the other hand, a client server system has a fixed set of servers, additional nodes or clients usually slow down the system and impact negatively on data transfer. P2P networks are distributed thereby increasing their robustness with the ability to replicate data over multiple peers. Thus, by enabling peers to find the data without relying on a centralized index in P2P systems, there is no single point of failure in the system [5] . For instance, in the Napsterization case, P2P was employed to empower nodes in association with a central index, resulting in speed and efficiency that translate to easy access to available content [6].

II. RELATED WORKS

The majority of security concerns related to instant messaging revolve around web based instant messaging or instant messages being exchanged over the Internet. These messages can easily be intercepted, "spoofed" or modified. This is of major concern to corporations who use this web based instant messaging to exchange confidential or business specific messages.

These security risks are however irrelevant to people using IMs for social reasons. Businesses however, are very conscious of the security risks associated with using web based instant messaging and as a result many have moved to instant messaging for local area networks [7]

Criminals engage P2P networks to also aid online anonymity. They usually do not wish to be identified as publishers (sender), or recipient (receiver), of information. This is because, in most cases, the materials being distributed are illegal or incriminating. At times, the materials may be legal but socially deplored, embarrassing or against workplace ethics, religion, or social beliefs. In some cases, P2P users fear retribution, censorship and therefore prefer a platform that guarantees personal privacy by using systems that prevents tracking or data mining activities. In another vein, some organization employ P2P systems to distribute materials whose distribution are legal, but in high demand. For instance, anonymous P2P networks can be used for speedy downloads of new release of computer software or media.

A. Anonymous Instant Messaging Systems:

For instance, in AnImA, a P2P system, a proxy for all XMPP compatible Instant Messengers acts as a server for the messenger software and routes the messages through a Peer-to-Peer network with a high grade of anonymity[8]. Drac is a peer-to-peer communication system designed to make IM and VoIP traffic anonymous and unobservable. It achieves this goal by exposing the social connections of the users who make up the nodes of the peer-to-peer network. Drac gives away the identity of a user's friends to guarantee the unobservability of actual calls, while still providing anonymity when talking to trusted third parties. The authors note that “while anonymous online communication may conceal the content of conversations, information about the network addressing, the timing of the messages, and the volume of traffic often reveals as much as the hidden correspondence” Drac aims to preserve anonymity while also thwarting traffic analysis by using a peer-to-peer relay architecture that routes data through social networking connections [9]. TorChat is a peer to peer instant messenger with a completely decentralized design, built on top of Tor 's location hidden services. It provides extremely strong anonymity while being very easy to use without the need to install or configure anything. It was developed on Linux with cross platform usability.

The designers claimed that “nobody will be able to find out where the user is even if they are already observing users and sniff internet connection they will not be able to find out what information is being sent or received, to whom the information is being send or being received as well as the location of the contacts. [10]

B. Previous Works:

Previous work on anonymous P2P systems have focused on two major approaches, Onion Routing [11] and the Crowds [12] approach. In Onion Routing messages are routed through Onion Routers which are a dedicated set of servers. A message is first encrypted into a layered data structure, known as an onion, which contains cryptographic information for every hop on the path of the message. It is then sent off to the first Onion Router in the path which is able to decrypt the first layer of encryption of the message. After decryption this router knows the address of the next router to forward the message on to. Eventually the last router in the path receives the message, decrypts the last layer of the message, and forwards the message on to its intended final destination.

Crowds is the anonymous protocol where nodes randomly decide to either forward packets on to other random nodes in the network or deliver them to their destinations. Successive requests in Crowds follow the same path from one node to another as previous requests until the network signals for a periodic path reformation to occur which usually happens hourly. Crowds also uses a centralized server to provide admission control for the network [13]. Peer-to-peer (P2P) anonymous communication systems are vulnerable to free-riders, peers that use the system while providing little or no service to others and whose presence limits the strength of anonymity as well as the efficiency of the system. Free -riding can be addressed by building explicit incentive mechanisms into system protocols to promote two distinct aspects of cooperation among peers compliance with the protocol specification and the availability of peers to serve others [13][14].

III. RESEARCH DIRECTION

The fact that it is difficult or impossible to hide the fact that a P2P network is used means that they could be simply outlaw to prevent the free flow of information. These drawbacks do not apply to P2P systems used on a wireless mesh network. In this scenario, users do not sign up with ISPs to participate in such a network, and are only identifiable through their hardware. Next we propose a framework that can be used to tracks the behaviour of different individuals while making use of P2P networks for information sourcing, communication through messaging applications and chat programmes on wireless networks.

A. Current P2P Chat Use Case Scenario:

The Existing system is the popular chatting systems like yahoo messanger, windows live, myspace just to mention a few. The basic usecase of such systems is directed to instant messaging using socket programming. Fig. 1 below is the use case diagram describing what the primary actor in the system does.

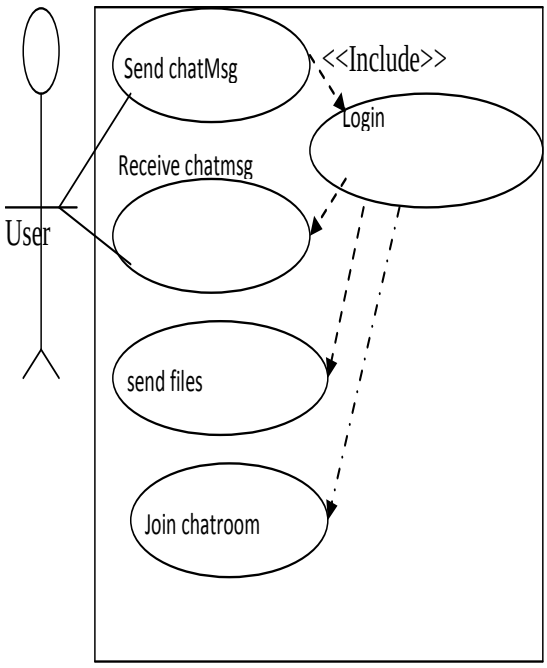

Figure. 1: Use Case Diagram

Fig 1depicts the basic operation performed on IM is sending and receiveing chat messages. These system are not designed in a way that the messages sent is checked for illicit content and therefore makes the system insecure. The activity diagram (Fig 2) depicts what happens when a user sends a chat message to a chat mate. The diagram assumes that the user is logged in. The user then selects the chat client he/she wants to chat with. If the chatmate is available, the message is displayed, otherwise the message is saved as an offline message.

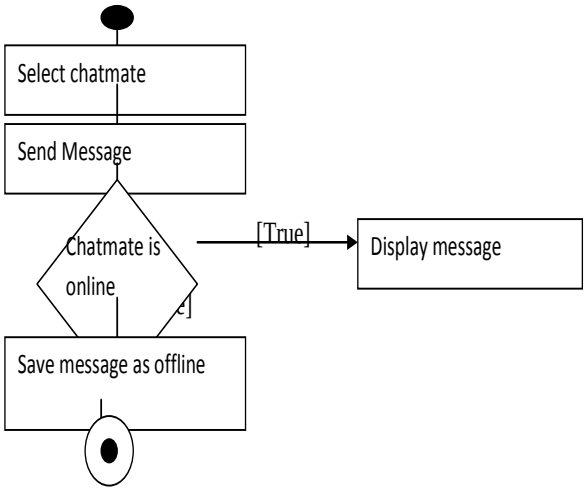

Figure. 2: Activity Diagram Sendchat Massage

B. Proposed System Architecture – Use Case:

Our proposed architecture differs from the existing existing system in that a secure component is integrated into it. We also provide a component that describes user behaviour for the system. Information in this security component is updated periodically each time users engage in chat sessions. The security module disallows illicit content from being propagated.

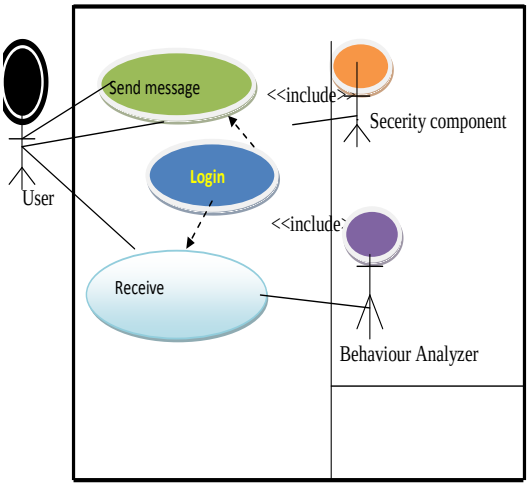

Figure. 3: Use Cases for Proposed System

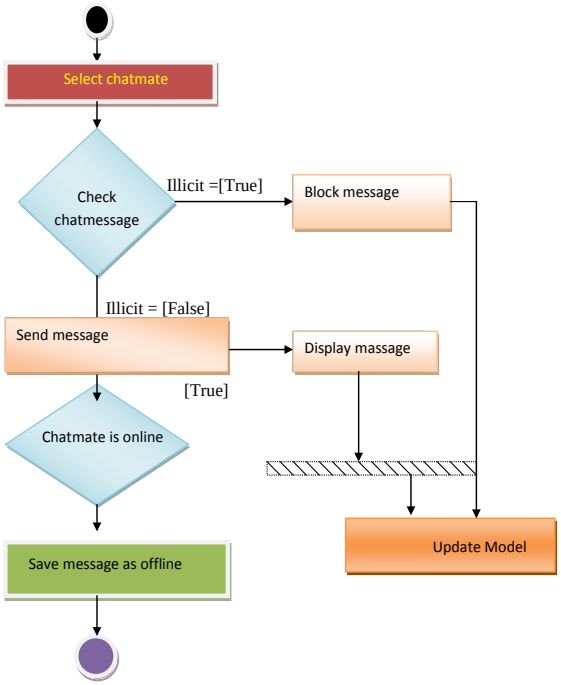

Figure. 4: Activity Diagrams Send Chat Message

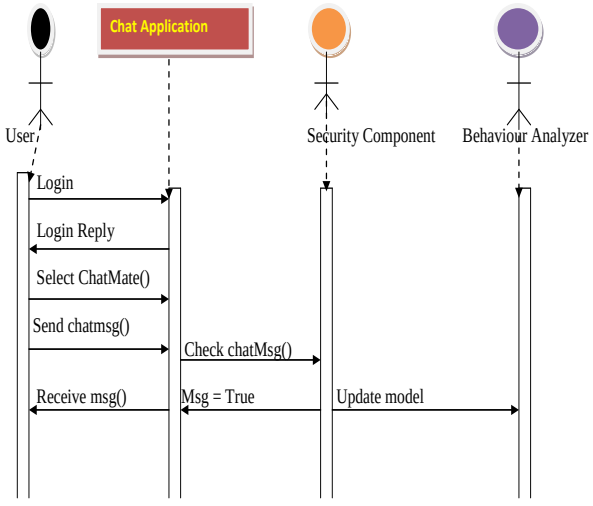

Figure. 5: Sequence Diagram (Send chat message)

Our architecture will enable the implementation of P2P Systems where usage and usage contents can be monitod using the security component . The work of the security component is to track the kind of messages the user send whenever the user logs in and send messages. It will also check the message for illicit contents (when true it blocks access) and when illicit is false it send message to chatmate online and display message and save message as offline and later update model. The role of the behaviour analyzer is to keep track of user interest while online and serve as a mechanism that recommends and suggest possible user offline behaviour.

IV. CONCLUDING REMARKS

We have proposed a framework that can be implemented to track user behaviour on a peer to peer messaging and chat system. The intention is to create a system-aware scenario for P2P networks such that P2P networks can become safer and by extension help human scientists in their quest to understand online user behaviour as a way of predicting offline interactions and conduct.

V. FUTURE WORKS

In the future, we hope to employ mathematical constructs to model user behaviour on the internet and by extension implement our proposed framework using appropriate software and hardware technologies. Object oriented paradigm will be used to create chart software and an internet explorer plug-in that monitors the behaviour of internet users.

VI. REFERENCES

[1]. S. Ramesh, & B. Goodman eds), 2005, Peer-to-Peer Computing: Evolution of a Disruptive Technology, Idea Group Inc., Hershey, PA, USA.

[2]. B. Lipinski & P. Mac Alpine (2006). A Security Review of an Anonymous Peer-to-Peer File Transfer Protocol. B. Lipinski and P. MacApline. A Security Review of an Anonymous Peer-to-Peer File Transfer Protocol. Technical report, Rice University, Houston, TX, USA.

[3]. O.B. Longe, O.B. S.C. Chiemeke, O.F.W. Onifade, O.F. F.M. Balogun, F.A. Longe & V.U. Otti (2007). Exposure of Children and Teenagers To Internet Pornography In South Western Nigeria – Concerns, Trends & Implications. Journal of Information Technology Impact. Vol 7, No. 3. Available online at

[4]. A. Rowstron & P. Druschel, 2001), Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems. In proceedings Middleware 2001: IFIP/ACM International Conference on Distributed Systems Platforms. Heidelberg, Germany. Lecture Notes in Computer Science, Vol 2218, Pg 329.

[5]. Steinmetz, R, Wehrle K (Eds) (2005). Peer-to-Peer Systems and Applications, Volume 3485, Sep 2005.

[6]. I. Stoica, R. Morris, D. Karger, M. Kaashoek & H. Balakrishnan (2001), a scalable peer-to-peer lookup service for internet applications. In Proceedings of SIGCOMM.

[7]. How Does Peer to Peer Instant Messaging Work

[8]. AnIMA http://anima.sourceforge.net/Anima.html

[9]. C. Thomas (2010). Secure P2P Scheme Leverages Social Networks. Anonymous and unobservable IM and VoIP could be possible under a proposed network architecture called Drac. InformationWeek http://www.informationweek.com/news/security/privacy/224400759

[10]. TorChat a anonymous Instant Messenger for Tor Network http://www.unitethecows.com/other-p2p-clients/46928-torchat-anonymous-instant-messenger-tor-network.html

[11]. Onion Routing. http://www.onion-router.net/

[12]. M. K. Reiter & A.D. Rubin (1999) Anonymous Web transactions with Crowds. Communications of the ACM, 42(2):32-48, February 1999.

[13]. M. Rennhard and B. Plattner (2002) Introducing MorphMix: Peer-to-peer based anonymous internet usage with collusion detection. Proceedings WPES, 2002, pp.91-102

[14]. D. Figueiredo, J. Shapiro, and D. Towsley (2005) "Incentives to promote availability in Peer-to-Peer anonymity systems", In Proceedings of IEEE ICNP, 2005. Daniel Figueiredo &

