# THE INFORMATION REVELATION AND PRIVACY IN ONLINE SOCIAL NETWORKS

**K.Govindaraj[1], Y.A.Sivaprasad[2]**

*[1,2]Dept of CSE, Chendhuran College of Engineering and Technology*

*Anna University, Chennai. (India)*

## ABSTRACT

*The client summary identical with privacy-preservation in movable community set of connections (msns) and begin a family of novel summary identical protocols. We first propose a clear Comparison-based summary identical protocol which runs between two parties, an originator and a responder. The summary identical protocol enables the originator to get hold of the comparison-based matching result about a specified attribute in their profiles, while avoid their characteristic values from disclosure. Then we propose an implicit Comparison-based Profile matching protocol which allows the originator to straight get hold of a number of messages instead of the comparison result from the responder. The messages not linked to user profile can be separated into many groups by the responder. The originator completely decides the interested group which is unidentified to the responder. Two messages in both groups are prepared by the responder, and only one message can be obtained by the initiator according to the comparison result on a single attribute. We further generalize the icpm to an implicit Predicate-based Profile matching protocol which allows complex comparison criteria spanning multiple attributes. The anonymity analysis shows all these protocols achieve the confidentiality of user profiles. In addition, the reveals the comparison result to the initiator and provides only conditional anonymity; they do not reveal the result at all and provide full secrecy. We analyze the communication overhead and the anonymity strength of the protocols. We then present an enhanced version of the called by combining the with a novel prediction-based adaptive pseudonym change strategy. The performance comparatively studied through extensive trace-based simulations. Simulation results demonstrate that they achieve significantly higher anonymity strength with slightly larger number of pseudonym.*

***Keywords: Mobile Social Networks (MSNS), Explicit Comparison-Based Profile Matching Protocol (ECPM), Implicit Comparison-Based Profile Matching Protocol (ICPM), Implicit Predicate-Based Profile Matching Protocol (IPPM)***

## I INTRODUCTION

Social networking makes digital communication technologies sharpening tools for extending the social circle of people. It has already become an important integral part of our daily lives, enabling us to contact our friends and families on time. Social networking sites such as Face book and Twitter have reached 82 percent of the world's

online population, representing 1.2 billion users around the world. In the meantime, fuelled by the pervasive adoption of advanced handheld devices and the ubiquitous connections of Bluetooth/Wi-Fi/GSM/LTE networks, the use of Mobile Social Networking (MSNs) has surged. In the MSNs, users are able to not only surf the Internet but also communicate with peers in close vicinity using short-range wireless communications. Due to its geographical nature, the MSNs support many promising and novel applications For example, through Bluetooth communications, People Net enables efficient information search among neighbouring mobile phones; a message-relay approach is suggested in to facilitate carpool and ride sharing in a local region. Realizing the potential benefits brought by the MSNs, recent research efforts have been put on how to improve the effectiveness and efficiency of the communications among the MSN users. They developed specialized data routing and forwarding protocols associated with the social features exhibited from the behaviour of users, such as, social friendship, social selfishness, and social morality. It is encouraging that the traditional solutions can be further extended to solve the MSN problems by considering the unique social features.

Privacy preservation is a significant research issue in social networking. Since more personalized information is shared with the public, violating the privacy of a target user become much easier. Research efforts have been put on identity presentation and privacy concerns in social networking sites. Gross and Acquits argued that users are putting themselves at risk both offline (e.g., stalking) and online (e.g., identity theft) based on a behaviour analysis of more than 4,000 students who have joined a popular social networking site. Stutsman presented a quantitative analysis of identity information disclosure in social network communities and subjective opinions from students regarding identity protection and information disclosure. When the social networking platforms are extended into the mobile environment, users require more extensive privacy-preservation because they are unfamiliar with the neighbours in close vicinity who may eavesdrop, store, and correlate their personal information at different time periods and locations. Once the personal information is correlated to the location information, the behaviour of users will be completely disclosed to the public.

Chen and Rahman surveyed various mobile Social Networking Applications (SNAs), such as, neighbourhood exploring applications, mobile-specific SNAs, and content-sharing applications, all of which provide no feedback or control mechanisms to users and may cause inappropriate location and identity information disclosure. To overcome the privacy violation in MSNs, many privacy enhancing techniques have been adopted into the MSN applications For example, when two users encounter in the MSNs, privacy-preserving profile matching acts as a critical initial step to help users, especially strangers, initialize conversation with each other in a distributed and privacy-preserving manner. Many research efforts on the privacy preserving profile matching have been carried out. The common goal of these works is to enable the handshake between two encountered users if both users satisfy each other's requirement while eliminating the unnecessary information disclosure if they are not. The original idea is from where an agent of the Central Intelligence Agency (CIA) wants to authenticate herself to a server, but does not want to reveal her CIA credentials unless the server is a genuine CIA outlet. In the meantime, the server does not want to reveal its CIA credentials to anyone but CIA agents. In the MSNs, we consider a generalized function to support information exchange by using profile matching as a metric. Following the previous example, we consider two CIA agents with two different priority levels in the CIA system, $A$ with a low priority $lA$ and $B$ with a high priority $lB$. They know each other as a CIA agent. However, they do not want to reveal their priority levels to each other. $B$ wants to share some messages to $A$. The messages are not related to user profile, and they are divided into multiple categories, e.g., the messages

related to different regions (New York or Beijing) in different years (2011 or 2012). *B* shares one message of a specified category *T* at a time. The category *T* is chosen by *A*, but the choice is unknown to *B*. For each category, *B* prepares two self-defined messages, e.g., a low-confidential message for the CIA agent at a lower level and a high-confidential message for the agent at a higher level. Because *lA* < *lB*, *A* eventually obtains the low-confidential message without knowing that it is a low confidential one. In the meantime, *B* does not know which message *A* receives. The above function offers both *A* and *B* the highest anonymity since neither the comparison result between *lA* and *lB* is disclosed to *A* or *B* nor the category *T* of *A*'s interest is disclosed to *B*. In the following, we refer to *A* as the initiator *ui*, *B* as the responder *uj,* the attribute used in the comparison (i.e., priority level) as *ax*, and the category *T* of *A*'s interest as *Ty*. The attribute values of *ui* and *uj* on the attribute *ax* are denoted by *ai; x* and *aj; x*, respectively. We first formally describe two scenarios from the above examples.

## II. RELATED WORK

The icpm and the ippm do not reveal the result at all and provide full anonymity. Users require more extensive privacy-preservation because they are unfamiliar with the neighbors in close vicinity who may eavesdrop, store, and correlate their personal information at different time periods and locations. The improved protocol only reveals whether the dot product is above or below a given threshold. The threshold value is selected by the user who initiates the profile matching. They pointed out the potential anonymity risk of their protocols. The threshold value must be larger than a pre-defined lower bound (a system parameter) to guarantee user anonymity. The homomorphic encryption schemes that support different operations such as addition and multiplication on cipher texts. The user is able to process the encrypted plaintext without knowing the secret keys. The dot product protocol is lack of verifiable secure computation. The Protocol only reveals whether the dot product is above or below a given threshold. They pointed out the potential anonymity risk of their protocols; an adversary may adaptively adjust the threshold value to quickly narrow down the value range of the victim profile. It Present an enhanced version of the ecpm, called ecpm+, by combining the ecpm with a novel prediction-based adaptive pseudonym change strategy. The performance of the ecpm and the ecpm+ are comparatively studied through extensive trace-based simulations. The ecpm+ achieves significantly higher anonymity strength with slightly larger number of pseudonyms than the ecpm. The msns, users are able to not only surf the Internet but also communicate with peers in close vicinity using short-range wireless communications. The social features exhibited from the behaviour of users, such as, social friendship social selfishness and social morality. It is encouraging that the traditional solutions can be further extended to solve the MSN problems by considering the unique social features. The homomorphic encryption schemes are widely used in data aggregation and computation specifically for privacy-sensitive information. We review the homomorphic encryption scheme that serves a building block of our proposed profile matching protocols. The profile matching protocols are novel since the comparison of attribute values is considered as the matching operation.

### 2.1 Efficient Private Matching And Set Intersection

This work considers several two-party set-intersection problems and presents corresponding secure protocols. Our protocols enable two parties that each holds a set of inputs. The set-intersection primitive is quite useful as it is extensively used in computations over databases, e.g., for data mining where the data is vertically

partitioned between parties. One could envision the usage of efficient set-intersection protocols for online recommendation services, online dating services, medical databases, and many other applications. We are already beginning to see the deployment of such applications using either trusted third parties or plain insecure communication. Malicious adversaries' fort their overhead for input lists of length k is O(k) communication and O(k log k) computation, with small constant factors. A simple reduction from the communication lower-bound on disjointness shows that this problem cannot have a sub linear worst-case communication overhead. We show a sampling-based private approximation protocol that achieves instance-optimal communication. In this model, an adversary may behave arbitrarily. In particular, we cannot hope to avoid parties (i) refusing to participate in the protocol, (ii) substituting an input with an arbitrary value, and (iii) prematurely aborting the protocol.

### 2.2 K-Anonymity: A Model For Protecting Privacy

Data holders, operating autonomously and with limited knowledge, are left with the difficulty of releasing information that does not compromise privacy, confidentiality or national interests. In many cases the survival of the database itself depends on the data holder's ability to produce anonymous data because not releasing such information at all may diminish the need for the data, while on the other hand, failing to provide proper protection within a release may create circumstances that harm the public or others. So a common practice is for organizations to release and receive person specific data with all explicit identifiers, such as name, address and telephone number, removed on the assumption that anonymity is maintained because the resulting data look anonymous. However, in most of these cases, the remaining data can be used to re-identify individuals by linking or matching the data to other data or by looking at unique characteristics found in the released data.

### 2.3 Homomorphic Encryption

The development of cloud storage and computing plat- forms allows users to outsource storage and computations on their data, and allows businesses to the task of maintaining data-centers. An excellent way to these privacy concerns is to store all data in the cloud encrypted, and perform computations on encrypted data. To this end, we need an encryption scheme that allows meaningful computation on encrypted data, namely a homomorphic encryption scheme. Homomorphic encryption schemes that allow simple computations on encrypted data have been known for a long time. We build on the somewhat homomorphic encryption, and implement simple statistics such as mean, standard deviation and logistical regression, and report on the performance number.

## III. PROPOSED APPROACH

It presents an enhanced version of the ecpm, called ecpm+, by combining the ecpm with a novel prediction-based adaptive pseudonym change strategy. The performance of the ecpm and the ecpm+ are comparatively studied through extensive trace-based simulations. The ecpm+ achieves significantly higher anonymity strength with slightly larger number of pseudonyms than the ecpm. The msns, users are able to not only surf the Internet but also communicate with peers in close vicinity using short-range wireless communications. The social features exhibited from the behaviour of users, such as, social friendship social selfishness and social morality it is encouraging that the traditional solutions can be further extended to solve the MSN problems by considering the unique social features. The homomorphic encryption schemes are widely used in data

aggregation and computation specifically for privacy-sensitive information. We review the homomorphic encryption scheme that serves a building block of our proposed profile matching protocols. The profile matching protocols are novel since the comparison of attribute values is considered as the matching operation.

## IV. SYSTEM DESIGN

### 4.1 System Architecture

The icpm for in this protocol, the responder prepares multiple categories of messages where two messages are generated for each category. The initiator can obtain only one message related to one category for each run. During the protocol, the responder is unable to know the category of the initiator's interest. To receive which message in the category is dependent on the comparison result on a specified attribute. The responder does not know which message the initiator receives, while the initiator cannot derive the comparison result from the received message. We provide an analysis of the effectiveness of the icpm, and show that the icpm achieves full anonymity.

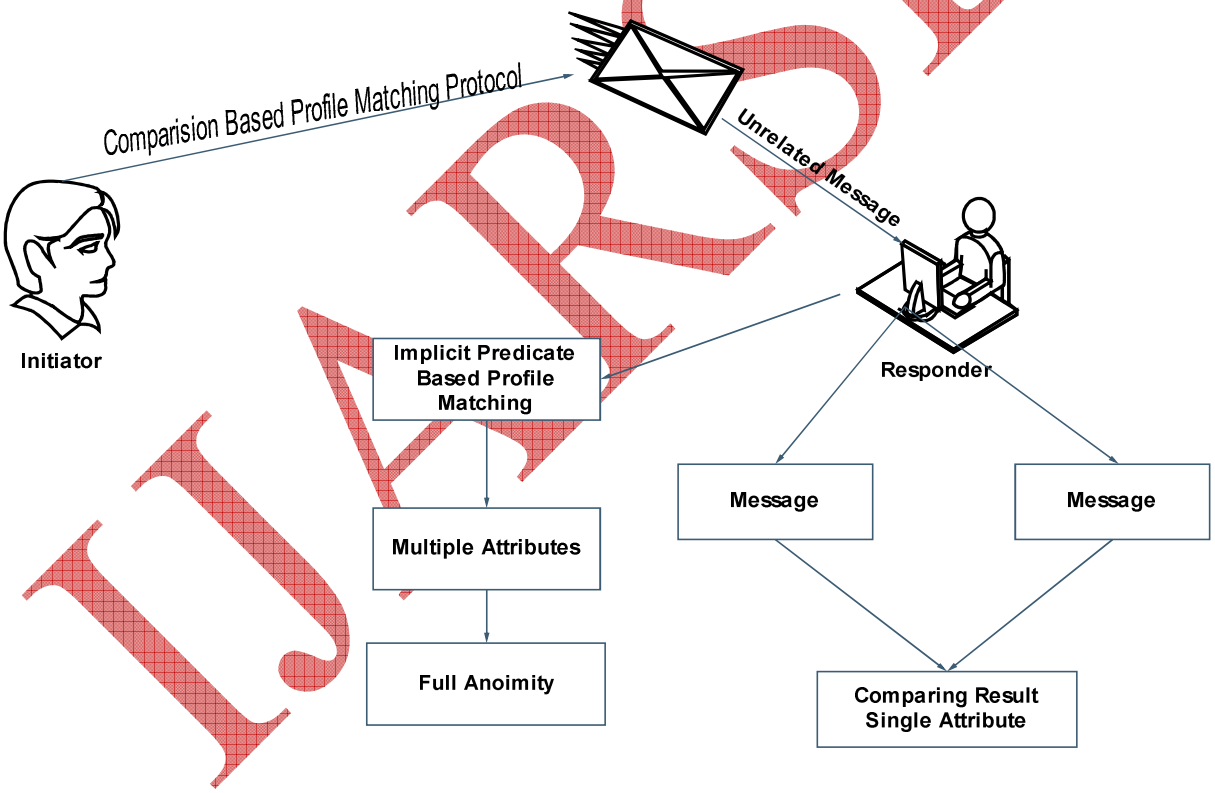


**Fig1.System Flow Diagram**

### 4.2 Modules

1. Profile Creator
2. Create Attribute Value
3. Initiator/Responder Secret Sharing
4. Implicit Comparison Based Profile Matching
5. Responder Matching Secrets

### 4.2.1 Profile Creator

The profile creator is the initial module to create an initiator profile or responder profile. The user profile details(username,password,email,mobile,gender,dob,occupation,address,location,hobbie/interest).

### 4.2.2 Create Attribute Value

An explicit Comparison-based Profile Matching protocol (ECPM) which runs between two parties, an initiator and a responder. The ecpm enables the initiator to obtain the comparison-based matching result about a specified attribute in their profiles, while preventing their attribute values from disclosure the attribute used in the comparison (i.e., priority level) as ax, and the category T of A's interest as Ty. The attribute values of on the attribute ax are denoted, respectively.

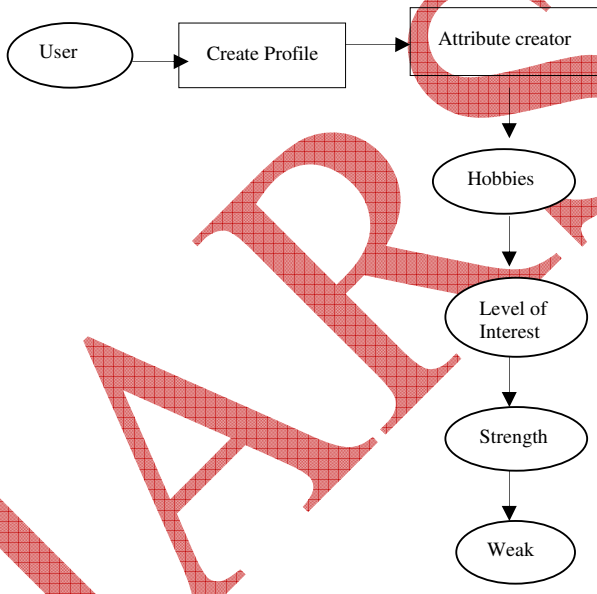


**Fig.2.Create Attribute Values**

### 4.2.3 Initiator/Responder Secret Sharing

The initiator expects that the responder shares one message related to the category of its interest, which is however kept unknown to the responder. In the meantime, the responder wants to share with the initiator one message which is determined by the comparison result of their attribute values. The comparison and the categories of messages the initiator first generates vector where the *y*-the dimension value is 1 and other dimension values are 0. Then, encrypts the vector with its own public key and sends the cipher texts to the responder. The cipher texts imply *ui*'s interested category *Ty*, but *uj* is unable to know *Ty* since *E*(0) and non-distinguishable without a decryption key.

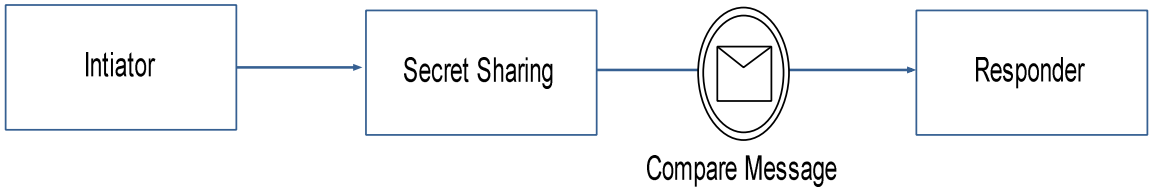


**Fig.3.Initiator/Responder Secret Sharing**

### 4.2.4 Implicit Comparison Based Profile Matching

The Implicit Comparison-based Profile Matching protocol (ICPM) which allows the initiator to directly obtain some messages instead of the comparison result from the responder. The messages unrelated to user profile can be divided into multiple categories by the responder. The initiator implicitly chooses the interested category which is unknown to the responder. Two messages in each category are prepared by the responder, and only one message can be obtained by the initiator according to the comparison result on a single attribute. Implicit Comparison-based Profile Matching (ICPM) and Implicit Predicate-based Profile Matching (IPPM). The icpm handles profile matching based on a single comparison of an attribute while the ippm is implemented with a logical expression made of multiple comparisons spanning multiple attributes. The icpm and ippm both enable users to anonymously request for messages and respond to the requests according to the profile matching result, without disclosing any profile information.

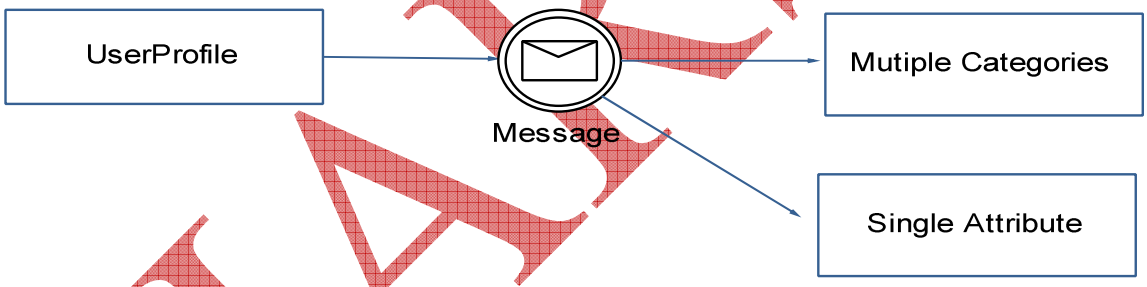


**Fig.4.Implicit Profile Matching**

### 4.2.5 Responder Matching Secrets

The initiator can obtain only one message related to one category for each run. During the protocol, the responder is unable to know the category of the initiator's interest. To receive which message in the category is dependent on the comparison result on a specified attribute. The responder does not know which message the initiator receives, while the initiator cannot derive the comparison result from the received message. We provide an analysis of the effectiveness of the iCPM, and show that the iCPM achieves full anonymity.

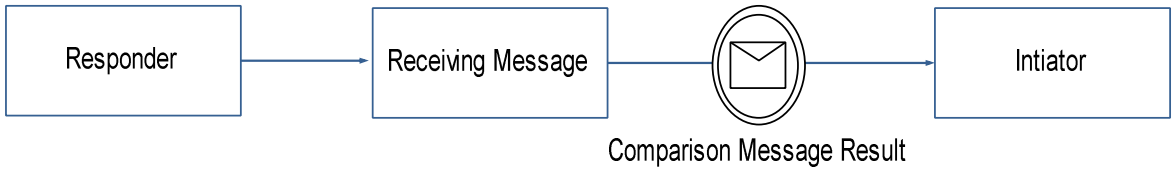


**Fig.5.Responder Matching Secrets**

## V. CONCLUSION

We have investigated a unique comparison-based profile matching problem in Mobile Social Networks (MSNs), and proposed novel protocols to solve it. The explicit Comparison based Profile Matching (ecpm) protocol provides conditional anonymity. It reveals the comparison result to the initiator. Consider the *k*-anonymity as a user requirement; we analyze the anonymity risk level in relation to the pseudonym change for consecutive ecpm runs. We have further introduced an enhanced version of the ecpm, i.e., ecpm+, by exploiting the prediction-based strategy and adopting the pre-adaptive pseudonym change. The effectiveness of the ecpm+ is validated through extensive simulations using real-trace data. We have also devised two protocols with full anonymity, i.e., implicit Comparison-based Profile Matching (icpm) and implicit Predicate-based Profile Matching (ippm). The icpm handles profile matching based on a single comparison of an attribute while the ippm is implemented with a logical expression made of multiple comparisons spanning multiple attributes. The icpm and the ippm both enable users to anonymously request for messages and respond to the requests according to the profile matching result, without disclosing any profile information. In current version of the icpm and the ippm, we implement “>” and “<” operations for profile matching. One future work is to extend them to support more operations, such as “≥” and “≤”. Another future work is to hide the predicate information in the ippm. Currently, the responder needs to transmit the threshold value of the predicate to the initiator, which may reveal partial information of the responder’s interest. Restricting the disclosure of such parameter will be of significance for advancing comparison-based family of profile matching protocols and warrants deep investigation.

## VI. FUTURE ENCHANCEMENTS

One future work is to extend them to support more operations, such as “≥” and “≤”. Another future work is to hide the predicate information in the ippm. Currently, the responder needs to transmit the threshold value of the predicate to the initiator, which may reveal partial information of the responder’s interest. Restricting the disclosure of such parameter will be of significance for advancing comparison-based family of profile matching protocols and warrants deep investigation.

## ACKNOWLEDGMENT

We would like to sincerely thank Assistant Prof. Y.A.Sivaprasad for his advice and guidance at the start of this article. His guidance has also been essential during some steps of this article and his quick invaluable insights have always been very helpful. His hard working and passion for research also has set an example that we would like to follow. We really appreciate his interest and enthusiasm during this article. Finally we thank the Editor in chief, the Associate Editor and anonymous Referees for their comments.

## REFERENCES

[1] A. G. Miklas, K. K. Gollu, K. K. W. Chan, S. Saroiu, P. K. Gummadi, and E. de Lara, “*Exploiting social interactions in mobile systems*,” in Ubicomp, 2007, pp. 409–428.

[2] S. Ioannidis, A. Chaintreau, and L. Massoulie, “*Optimal and scalable distribution of content updates over a mobile social network*,” in Proc.IEEE INFOCOM, 2009, pp. 1422–1430.

[3] R. Lu, X. Lin, and X. Shen, "*Spring: A social-based privacy-preserving packet forwarding protocol for vehicular delay tolerant networks*," in Proc. IEEE INFOCOM, 2010, pp. 632–640.

[4] W. He, Y. Huang, K. Nahrstedt, and B. Wu, "*Message propagation in adhoc-based proximity mobile social networks*," in PERCOM workshops, 2010, pp. 141–146.

[5] D. Niyato, P. Wang, W. Saad, and A. Hjørungnes, "*Controlled coalitional games for cooperative mobile social networks,*" IEEE Transactions on Vehicular Technology, vol. 60, no. 4, pp. 1812–1824, 2011

[6] M. Motani, V. Srinivasan, and P. Nuggehalli, "*Peoplenet: engineering a wireless virtual social network*," in MobiCom, 2005, pp. 243–257.

[7] M. Brereton, P. Roe, M. Foth, J. M. Bunker, and L. Buys, "*Designing participation in agile ridesharing with mobile social software*," in OZCHI, 2009, pp. 257–260.

[8] E.Bulut and B.Szymanski, "*Exploiting friendship relations for efficient routing in delay tolerant mobile social networks,"* IEEE Transactions on Parallel and Distributed Systems, vol. 23, no. 12, pp. 2254–2265, 2012.

[9] Z. Yang, B. Zhang, J. Dai, A. C. Champion, D. Xuan, and D. Li,"*E-smalltalker: A distributed mobile system for social networking in physical proximity*," in ICDCS, 2010, pp. 468–477.

[10] Q. Li, S. Zhu, and G. Cao, "*Routing in socially selfish delay tolerant networks,"* in Proc. IEEE INFOCOM, 2010, pp. 857–865.

[11] X. Liang, X. Li, T. H. Luan, R. Lu, X. Lin, and X. Shen, "*Morality driven data forwarding with privacy preservation in mobile social networks,*" IEEE Transactions on Vehicular Technology, vol. 7, no. 61,pp. 3209–3222, 2012.

## AUTHORS PROFILE

**K.Govindaraj** is currently a PG Scholar in Computer Science Engineering from the Department of Computer Science and Engineering at Chendhuran College of Engineering and Technology, Pudukkottai. He received his Bachelor Degree in Computer Science from Bharathidasan University Trichy and Tamilnadu. His Research areas include Data Mining, data warehousing and distributed system.

**Y.A.Sivaprasad** is currently working as an Associate Prof. from the Department of Computer Science and Engineering at Chendhuran College of Engineering and Technology, Pudukkottai. He Published 3 Conferences and 5 international journals. His main research interests lie in the area of Data Mining, wireless sensor networks and distributed computing.