IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 16, 2019, accepted January 19, 2020, date of publication February 3, 2020, date of current version February 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2971239

Privacy Preservation Using 2-Degree Anonymity With Trust Circle in Ubiquitous Network for Service Communications

SWARNALI HAZRA AND S. K. SETUA
Department of Computer Science, University of Calcutta, Kolkata 700073, India
Corresponding author: Swarnali Hazra (swarnali.hazra@gmail.com)

ABSTRACT The ubiquitous networks bring lots of convenience to consumers and service providers for their service communications but have strong privacy issues. Such network environments are unobtrusive and imperceptible to the consumer, where services are provided through an omnipresent way in smart environments. It raises numerous privacy and security issues for consumers. Privacy of service communicating entities needs to be preserved from malicious entities in the context of three-fold privacy threat - identity, location, behavioural privacy threat of legitimate entities in ubiquitous environment. In this paper, privacy preservation is explored with two levels of anonymization by a 2-degree anonymity approach through trust circle of one service entity in service communication. Our proposal provides the personalization in-between anonymity and trust. Simulation results exhibit the effectiveness of our proposal in the considered environment with the "ADULT" database and "Facebook" data set of two Universities, Amherst, and Colgate.

INDEX TERMS Anonymity, dummy message, privacy, service entity, trust, trust circle.

I. INTRODUCTION

Ubiquitous networks [39] has been integrated with our daily life due to its dynamic nature and easily communicable platform. Such network environments provide service facilities to consumers in an omnipresent way through embedded computation and communication. Ubiquitous network supports mobility, imperceptibility, embedded smart environment, localized scalability, and uneven conditioning. In a ubiquitous environment, service providers provide services to consumers through an embedded platform based on consumer's service requirements in a distributed and imperceptible way. Embedded nature provides the unassuming environment to the consumers.

In a ubiquitous network environment [39], one service is served by the service provider after service discovery [38] that initiated by consumers, and sometimes initiated by the advertisement of service providers. Some malicious entities pretend themselves as legitimate service entities by intercepting the identity of target one and demands private information of its consumers to exploit. Sometimes the location of a legitimate service entity may be detected from location information of service packet, or by backtracking the path of the packet flow, or by following the packet flow direction. That location information can be misused in terms of authenticity. With the continuous observation of communications, some malicious entities may identify a legitimate entity from its behavioural characteristics and may misuse private information. This three-fold privacy threat to identity, location, and behavioural information necessitates a strong level of privacy barrier for consumers and service providers during service communication.

Privacy is determined by the allowance of one entity, about what information should be revealed from where and how. For preserving privacy from malicious activity, two levels of anonymization are introduced by our proposed 2-degree anonymity through the trust circle of every service entity in ubiquitous networks. Every service entity builds their trust circle depending on context-based interactions and direct-indirect trust evaluations. The trust [9], [37], [49]–[53] relation with entities is expanded from single hop to multi-hop by maintaining progressive trust [8], [13], [16], [24]. The scenario, where multi-hop distant nodes are not in direct interaction, the belief level may decay with the increase of hop-count. In this scenario, we have introduced progressive trust. 1st degree anonymity

The associate editor coordinating the review of this manuscript and approving it for publication was Meng-Lin Ku.

is introduced with trust-based (p, β) sensitive k-anonymity with alias originator (AlO). Here the identities of k number of trusted entities, having similar non-sensitive attribute values are reported as a source in service packet to introduce anonymity. These k trusted entities have at least p number of different sensitive attribute values and β number of sensitive attribute types of a relevant context. Actual originator (AcO) chooses (p, β) sensitive k-anonymity satisfied trusted entities from its trust circle, and selects an AlO from them. AcO sends a service packet to AlO. AlO routes this packet acting like an actual packet originator. On the other hand, 2^{nd} degree anonymity is introduced by the n-deviation with dummy messages from every intermediate trusted entity between AcO and AlO to deviate malicious entity. Here, 'n' is depended on the flexible ambiguity factor. In our 2-degree anonymity approach, AlO is chosen through a personalization between anonymity and trust [9], [37], [49]–[53].

The structure of the paper is as follows: Related works are discussed in section II. In section III, symbols, motivation, contributions, definitions, and roles of stakeholders are discussed. A three-fold privacy threat model is explained briefly in section IV. Network model, Objectives, and Three-fold Privacy Model are presented in section V. In section VI, the process of trust circle establishment is detailed out. The “three-fold privacy preservation with 2-degree anonymity” is described in section VII. The computational complexity of our proposal is discussed in section VIII. Section IX explains simulation results, and section X concludes the proposed work.

II. RELATED WORK

Nowadays, privacy preservation in ubiquitous networks is a challenge. During service packet transfer between two service entities in service communication, one's privacy can be forged by an adversary to misuse it.

Many privacy preservation approaches are proposed for different environments. The cryptographic approach is one of the traditional privacy preservation techniques. In [3], Gajparia introduces his proxy-based approach with the concept of trusted third party LIPA (Location Information Performance Authority). LIPA takes decisions of location information distribution based on user-chosen constraints. In [35], a two-party key agreement scheme is introduced with unilateral privacy based on pairing. In [19], ‘Chaums Blind Signature’ is used to achieve anonymity and privacy. In [32], the public key encryption process is used with the concept of a trusted computing module. In [40], User's privacy is preserved by a telecommunication service provider (TSP) through the exchange of encrypted feedback scores to the public bulletin board for other TSPs instead of private information exchange. TSP evaluate the reputation of its user from call record information and publishes the encrypted reputation score to the public bulletin board. The collaborating TSP or the protocol initiator then homomorphically computes the aggregated weighted average score for the user from the encrypted scores without learning scores provided by the collaborating TSP. In [42], when the consumer receives the requested product, he/she is asked for feedback about the seller. The consumer provides feedback to the bulletin board in an encrypted form for future weighted average computation by others. But the encryption-decryption process contradicts the openness and distributed nature of the ubiquitous environment and desired responsiveness of intended service communications.

Nowadays, anonymity is introduced to preserve privacy. In [27], the authors proposed a location privacy scheme through caching and special k-anonymity. They introduced a caching at anonymizer that maintains previous results of the user's point of interest (POI). If the requested location is available in the cache, users get it directly from anonymizer without any direct interaction with unreliable Provider (LSP). Otherwise, the anonymizer initiates special k-anonymity, where k numbers of location points are selected as per users' possible location predicted by the Markov method. But this approach is only applicable for users, not for provider privacy due to the cache overhead of numerous users' information. In [41], authors introduced a centralized global reputation system that computes the global reputation of the caller by weighted aggregation of the local reputation scores provided by the respective collaborating SPs without compromising the privacy. They strip off some private information of call record and uses k anonymized out-degree along with pseudo-identity of users to preserve the privacy. In [43], a User Equipment (UE) is registered to the network with its actual identity in the 5G roaming environment. In the handover phase, when a UE moves from source AP to target AP, mutual authentication with the key agreement between the UE and target AP proceeds based on three short message exchanges (request, reply, and acknowledgment). Authentication is done by using chameleon hash functions. In this approach, user equipment's identity-hiding and anonymity are achieved using pseudo-identity instead of the actual in the handover phase.

In ubiquitous networks, mobile nodes (MN) get their requisite services in foreign networks that are enabled by the extended services of their home network. A mutual authentication scheme is introduced in Shin *et al.* and Wen *et al.* [47], [48] for the roamed mobile node and foreign network through the home network of the mobile node. Both claimed that their schemes satisfy all the security requirements for the ubiquitous system. However, Farash *et al.* [46] showed that these schemes are still vulnerable to several attacks and proposed an improved authentication scheme with anonymity for consumers, roaming in ubiquitous networks. Here, MN's identity contains a nonce for randomization and is well protected by the secret key. Therefore, this message is fresh in each session such that it is infeasible for an adversary to identify the MN's identity or link any two sessions of the same MN. Choudhury *et al.* [45] showed that Farash *et al.* [46] is still vulnerable to several attacks and proposed an improved scheme i.e. privacy-preserving password-based authentication scheme for roaming in ubiquitous networks to solve the security issues of Farash *et al.*'s scheme [46]. However, Chaudhry *et al.*'s scheme [45] is still vulnerable to the Stolen-mobile device, the impersonate attack, etc. An improved and anonymous biometric-based authentication technique is introduced to overcome these security issues in [44] for roaming in ubiquitous networks. Mobile node registers with the home agents using identity, password, and biometrics. The mobile node and foreign agent perform mutual authentication and share the session key with the support of the mobile node's home agent. Bio-hash is applied in the implementation of a three-factor authentication (identity, password, and biometrics) in biometric-based authentication. The pseudo-identity is introduced for the identity of the mobile node after the authentication by the home agent to preserve privacy in future communication.

Nowadays, the most popular anonymity approach is being used by introducing k-anonymity to preserve privacy. According to Pfitzmann and Koehntopp, anonymity is introduced as a nonidentifiable state, based on a set of subjects [22]. In k-anonymity, every subject in a set is indistinguishable from at least other (k − 1) subjects. In [15], Sweeni raised the k-anonymity concept with the assumption of a quasi-identifier (QI). It is a set of attributes that may serve as an identifier in the data set where each tuple refers to an individual. A data set is considered as k-anonymity satisfied if quasi-identifier appears with at least k occurrences for individuals in the data set. In [14], Minimal Generalization Algorithm (MinGen) is proposed to provide k-anonymity protection with minimal distortion combining generalization and suppression techniques. Generalization replaces a value with a reduced form of precise value but semantically consistent. Suppression does not release value at all. In [17], k-anonymity is introduced with the concept of middleware architecture. Initially, nodes establish an authenticated and encrypted connection with the central anonymity server. When a mobile node communicates with external service, the anonymity server provides k-anonymity with spatial or temporal cloaking to achieve desired anonymity. In Spatial cloaking, a 2D point location of a mobile client is replaced by a spatial of (k-1) other mobile clients, where the original 2D point lies anywhere within the range. Temporal cloaking is the replacement of a time point with a time interval where the original time point lies. In [6], a flexible personalized privacy framework is introduced with k-anonymity for a wide range of mobile clients to achieve context-sensitive privacy. This framework supports the minimum level of desired anonymity and maximum temporal-spatial tolerances of acceptance willingness. An improved k-value is proposed in [7] to overcome the difficulties of choosing suitable k in personalized k-anonymity approaches. This model connects the user and location-based service provider, based on the trusted third-party model. In [25], an enhanced k-anonymity approach: (α,k)-anonymity model is proposed for privacy-preserving data publishing. Here privacy protection of relationship to sensitive attribute values is provided along with privacy protection of individual identity. α is a fractional value, and k is an integer value. After satisfying the k-anonymity, the frequency of sensitive attribute value is restricted within α value. In [28], a query initiator reports the coordinates of a rectangle with other (k-1) agents instead of its exact coordinate to provide k-anonymity in an ad-hoc network. Additionally, an anonymous selection algorithm considers a query requester that acts on behalf of the query initiator. Query requester is being selected from the agents in the k-1 rectangle.

In [4], Machanavajjhala shows k-anonymity can not protect the privacy against background knowledge attacks if there are insufficient diversity insensitive attributes. With this motivation, Machanavajjhala introduces a l-diversity approach that endorses the intra-group diversity with at least l well-represented sensitive values in anonymization mechanism where (l-1) damaging pieces of background knowledge are needed. In [5], the conventional k-anonymity model and l-diversity model are extended from relational data to social network data with the concept of k-anonymity and l-diversity approaches. In [23], based on the advantage of k-anonymity and l-diversity, the l-diversity concept is used in k-anonymity with the external data set. In [12], considering the individual's privacy requirement, the PKDLD (personalized k-degree-l-diversity) anonymity model based on the KDLD anonymity model is proposed. Here individuals can specify whether others can access their own friends' information and sensitive attributes or not. Three types of privacy attributes for individuals are introduced to develop a personalized k-degree-l-diversity (PKDLD) anonymity, model.

In [30], p-sensitivity is considered with k-anonymity in contrast to l-diversity. It requires k-anonymity constraint, and the distinct values for each sensitive attribute occur at least p times within the same group. In [34], (p, α) sensitive k-anonymity and (p+, α) sensitive k-anonymity is introduced. (p, α)sensitive k-anonymity deals with p different values for each sensitive attribute in each similar quasi-identifiers group with at least a total weight of α. (p+, α) sensitive k-anonymity deals with the p numbers of distinct categories for each sensitive attribute with at least α total weight. In [2], the authors specified quasi-identifiers' generalization constraints and introduced p-sensitive k-anonymity within imposed constraints. In [33], p+ sensitive k-anonymity and (p, α) sensitive k-anonymity is used. In [18], a three-stage algorithm is proposed with the concept of l-diversity, p sensitive k-anonymity, p+ sensitive k-anonymity, and t-closeness. Here, three stages are attribute weighting, data processing, and data anonymization. Some privacy-preserving scheme [10], [11] introduces path diversion by creating dummy paths using dummy message flow to preserve the location privacy. During the traversal of messages from actual source to destination through the actual path, intermediate nodes create dummy paths by flowing dummy messages. The path diversion concept deviates the malicious entity from tracing the actual path and final destination.

TABLE 1. Comaring the pros and cons.

Algo	Cons	Pros
[32, 35, 40, 42]	1. Encryption-decryption process contradicts the openness and distributed nature of ubiquitous environment, and desired responsiveness of intended service communications.	Supports privacy, maintains confidentiality and authentication.
[44, 45]	1. Difficult to find suitable hash function in an open environment where entities are free to join and leave, and get services from anywhere at any time. 2. Pseudo identity is not enough for anonymity to prevent privacy.	Supports anonymity, privacy, confidentiality, and authentication.
[15]	1. k-anonymity can not protect the privacy against background knowledge attacks if there are insufficient diversity insensitive attributes. [4,34]. 2. If any member among k number of members is compromised, privacy will be compromised.	Supports the openness of ubiquitous nature of environment.
[4]	1. Difficult to achieve well represented l number of sensitive values with diversity in each combination of QI group. [34] 2. Prone to skewness attack, similarity attack. [34] 3. Results greater distortion . 4. If any member among k number of members compromised, privacy will be compromised.	Overcome the problem in k-anonymity and protect against attribute disclosure by improved k anonymity with l diversity.
[30]	1. Sensitivity of various values under a sensitive attribute may similar which permits the information to be disclosed and large data utility loss.[18]./ 2. Results greater distortion . 3. If any member among k number of members is compromised, privacy will be compromised.	Overcome the difficulties of implementing l well represented values to protect against attribute disclosure by improved k anonymity with p sensitivity.
[33, 34]	1. Difficult to achieve different values of sensitive attribute with a total weight of α, and different categories in ubiquitous network. 2. Results greater distortion . 3. If any member among k number of members is compromised, privacy will be compromised.	Overcome the problem of information disclosure and utility loss in p sensitive k anonymity approach.

In Table 1, we have compared the pros and cons of our related work. In comparison to previous approaches, our proposed approach overcomes the mentioned problems. Our approach supports three-fold privacy: identity, location, behavioural privacy. It supports identical sensitive attribute values under a single sensitive attribute type to protect the attribute disclosure and results in less distortion.

III. SYMBOLS, MOTIVATION, CONTRIBUTIONS, AND DEFINITION

In this section, we have described the symbols used in our proposed scheme and our motivation towards this work. The roles of considered networking entities are described following the definitions related to our work.

A. SYMBOLS

Symbols of our proposed scheme is described in Table 2.

TABLE 2. Symbols used in our proposed scheme.

Symbol	Description
CN	Consumer.
SP	Service Provider.
SE	Service Entity. Consumers and service providers are considered as service entity.
TR	Trustor.
TE	Trustee.
R	Recommender.
DTE	Direct Trustee.
RET	Recommended Trustee.
AcO	Actual Originator of packet.
AlO	Alias Originator of packet.
DT	Direct Trust.
IT	Indirect Trust.
RT	Recommended Trust.
FT	Final Trust.
$P_{leg}(true\|p)$	Probability of being legitimate with the true value of the interaction.
$[_A(T)_B]^t$	A's trust (T) on B for trust computing time t. T can be replaced by DT, IT, RT, FT.
TC_{TP}	Trust Circle, a group of trusted entities.
$Intr^c_q$	q^{th} interaction based on context c.
C	Candidate set that contains all sensitive and non sensitive attributes as record of each TC_{TP} member.
NA_j	j^{th} non-sensitive attribute in C.
Rcd_i	i^{th} record in C.
NAv_j	j^{th} non-sensitive attribute value in Rcd_i.
QI	Quasi Identifier with the set of non sensitive attributes, where QI =$\{NA_1, NA_2, \ldots .NA_r\}$.
$S(NAv)_i$	i^{th} set of QI values, where $S(NAv)_i = \{NA_1, NA_2, \ldots .NA_r\}$.
$O_{S(NAv)_i}$	Number of occurrence of (S(NAv)i) in all records of a single QI group.
$p_{(QI)}$	Distinct number of sensitive attribute values, associated with all records in a single QI group .
$\beta_{(QI)}$	Distinct number of sensitive attribute types,associated with all records in a single QI group.
H_{cur}	Current height variable.
$H\left[(NA_j)_{Rcd_i}\right]$	Height of j^{th} non sensitive attribute of Rcd_i.
tn	Total number of context dependent interactions of SE with one DTE.
pn	Total numbers of positive interaction.
nn	Total numbers of negative interactions, that can be represented by (tn-pn).
Im_i	i^{th} intermediate trusted node that lies in actual packet flow path between AlO and AcO.
N_{nbr}	One entity's total numbers of neighbours.
n	One's target number of neighbours to which dummy messages will be sent to create dummy path.

B. MOTIVATION

In ubiquitous environments, disclosure of private information is a major concern of security threats as of now. The centralized approaches to enforce security for the ubiquitous environment contradict the openness of the stakeholders. In contrast to the traditional centralized approach, anonymity-based privacy preservation approaches motivated us to provide 2-degree anonymity in service communication through one's trust circle for preserving privacy. The 1^{st} degree anonymity, "(p, β) sensitive k-anonymity with AlO", is influenced by the different modified approach of k-anonymity to preserve the three-fold privacy in terms of one's identity, location information, and behavioural patterns. On the other hand, 2^{nd} degree anonymity, 'n-deviation with dummy message', is influenced by different path diversion concept to misguide malicious entity in tracing actual path. The problem of tracing one entity by malicious entity (following message flow direction or backtracking) motivated us to incorporate a mechanism of dummy message flow that will deviate malicious entity from the actual path to reach to a legitimate entity.

In society, people normally form their communities with others based on trust relationships [9], [37], [49]–[53], and this has motivated us to focus on building up one's trust circle that participates in 2-degree anonymity. The identity hiding concept motivated us not to reveal the actual identity, and to follow the attributes matching [16], [20], [26] of entities in building or extending one's desired trust circle. Since entities are not in direct interaction in multi-hop non-interactive trust relations, the belief may affect with hop-count increase. Indirect interactions in multi-hop trust evaluation motivated us to maintain the hop-by-hop progressive trust [8], [13], [16], [24], so belief level of SE does not decay down for its non-interactive recommended entities with the increase of hop-count between them. Our context-dependent direct and indirect trust evaluation is used to create, expand, and modify their trust circle concerning time.

C. CONTRIBUTIONS

The main contributions of our work are summarized as follows:

1) Formation of service-based groups of trusted entities for each stakeholder of the network, including the directly trusted entities and more belief sensitive recommended entities.
2) Creation of dynamic alias originator through a certain quantified level of anonymization and personalization of associated attributes and in place of actual service requester or service provider from the trusted group of entities of each stakeholder for service communications.
3) Creation of suspicious paths to misguide the adversaries.
4) Computation of trust values of the entities to block and(or) to make thoroughfare of service communication messages.
5) Simulation with readily available data sets.

D. DEFINITIONS

Some definitions related to our proposed approach are described as follows:

- *Trust:*
  For us, trust is defined as the measure of one's belief level to others based on actions in interactions along with the recommendations. According to Diego Gambetta [37], "Trust is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both before he can monitor such action (or independently or his capacity ever be able to monitor it) and in a context in which it affects his own action." If trust value denoted by T and probability of being legitimate with the true value of the interaction is denoted by $P_{leg}(true|p)$ then, the trust relationship is described as per "(1)". Here p is probability value.

$$T = \begin{cases} \textit{highest trust} & For, 0.9 < P_{leg}(true|p) \leq 1.0 \\ \textit{trust} & For, 0.5 < P_{leg}(true|p) \leq 0.9 \\ \textit{uncertain trust} & For, P_{leg}(true|p) = 0.5 \\ \textit{distrust} & For, 0.1 \leq P_{leg}(true|p) < 0.5 \\ \textit{highest distrust} & For, 0.0 \leq P_{leg}(true|p) < 0.1 \end{cases} \tag{1}$$

- *Service Entity (SE):*
  Consumer and service provider, between whom service communication is done, is considered as service entities.
- *Trustor (TR):*
  The entity which evaluates the trust level of another entity is considered as trustor. A service entity (consumer or service provider) is defined as a trustor when it interacts with another entity to get back the reply or reaction for measuring the trust level of that entity.
- *Trustee (TE):*
  The entity, whose trust level is being evaluated by the trustor, is considered as trustee.
- *Direct Trustee (DTE):*
  The entity who is in direct communication of TR and whose trust level is being evaluated directly by TR, is called direct trustee (DTE) of TR.
- *Recommended Trustee (RTE):*
  The entity, who is not in direct communication with TR but recommended to TR by a trusted entity from TR's trust circle for trust evaluation, is considered as the recommended trustee (RTE) of TR, based on the progressive trust acceptance.
- *Direct Trust (DT):*
  Direct trust is the measure of trust level on DTE, which is directly evaluated by TR based on context-dependent direct interactions.
- *Indirect Trust (IT):*
  Indirect trust is the cumulative effect of context-dependent recommendations.
- *Trust Circle* (TC_{TP})*:* SE's group of trusted entities, to whom it can trust for communication in the network, is called trust circle. The trust circle may expand from single-hop to multi-hop. Considering "(2)", A's trust circle (TC_{TP}) consists of trusted entities B, C, D, and E. The notation "→" denotes trust relation. Here A trusts B, B trusts C, C trusts D, and D trusts E. Consequently, A trusts C, D, and E based on the respective recommendations. These trust dependencies are not symmetric between trusted entities. A trusted entity may leave the TC_{TP} by reporting SE. If the leaving trusted entity is an intermediate recommender of other trusted member(s), SE will recompute the trust of those through the recommendation of other trusted members about them. SE updates that information in TC_{TP} for future

selection of '(p, β) sensitive k-anonymity' group. If the other recommendations about those trusted entities are not available, they are no longer included in TC_{TP} until new recommendations come. On arrival of new entities, TC_{TP} will be updated through trust evaluation of them as DTE or RTE.

$$\begin{aligned} for,\ & A \rightarrow B(interactive) \rightarrow \\ & C(non-interactive) \rightarrow \\ & D(non-interactive) \rightarrow \\ & E(non-interactive) \\ & A \rightarrow C,\ A \rightarrow D,\ A \rightarrow E \end{aligned} \tag{2}$$

- *Progressive Trust:*
  In a multi-hop trust relation, the increasing trust level with respect to every intermediate node is considered as progressive trust. The belief of the recommended entity may decay with the increase of hop-count, where TC_{TP} establishing SE is not able to interact with the recommended non-interactive multi-hop distant entities. Here in "(2)", A is not able to interact with the C, that recommended by B (trusted group member of the A's TC_{TP}). If (trust of B to C) $\geq$ (trust of A to B), C will be evaluated by A. If the evaluated trust of C is greater than 0.5, it will be included as a trusted member of A's TC_{TP}. When non-interactive C becomes member of A's TC_{TP}, it can recommend its own trusted members to A. When C recommends its trusted D to A via B (see "(2)"), the entity C and D both are non-interactive to A. In this way, non-interactive E is recommended to A via non-interactive D, C and interactive B. With the increase of such hop-count, non-interactive recommendation dependencies increases and the belief level of the TC_{TP} establisher on non-interactive multi-hop distant entities may decay. In this scenario, progressive trust [8], [13], [16], [24] is introduced where E is considered as A's RTE and being evaluated only when (trust of D to E) $\geq$ (trust of A to D). Computation overhead is minimized through the acceptance of recommended entities based on progressive trust as it only accepts and computes the qualified non-interactive trust chain relation.

E. ROLES OF NETWORK ENTITIES

Some roles of network entities related to our work are described as follows:

- *Consumer (CN):*
  It is an SE that requests its required service with a service type, description, related attributes, and relevant context from available service providers.
- *Service Provider (SP):*
  It is an SE that provides the consumer with the required services.
- *Trust circle (TC_{TP}):*
  Help to preserve SE's privacy by participating in anonymity during service communication.
- *Recommender (R):*
  Gives recommendations about its trusted entity to a SE for extending that SE's TC_{TP} on judging the context and attributes.

IV. THREE-FOLD PRIVACY THREAT MODEL

Service communication in the pervasive environment may involve threats to identity privacy, location privacy, and behavioural privacy. The three-fold privacy threat is discussed as follows.

A. IDENTITY PRIVACY THREAT

Stealing of one entity's identity information such as ID, name, date of birth, etc., is considered as an identity privacy threat. In a ubiquitous environment, service communication is done between service entities with different types of service packets (service request, service response, service advertisement, service, etc.). Service packets contain the identity of the service entity. A malicious entity collects the identity-related information from service packets. Malicious entity constructs identity portfolio with the collected identity-related information to identify the actual service entity.

In Fig. 1, we have shown a network segment having one malicious entity, M. It is collecting the identity-related information of consumer CN-1 and service provider SP-1. Here, CN-1 and SP-1 forwarded service packets contain their identity-related information. CN-1 forwarded service packet is received by SP2, SP3, and interpreted by the M. On the other hand, SP-1 forwarded service packet is received by CN-2, CN-3, and M. M constructs "Identity Portfolio" of CN-1 and SP-1 from their stolen identity-related information. As a consequence, M may identify the actual identity of the target by creating their actual profile from this "Identity Portfolio".

B. LOCATION PRIVACY THREAT

A malicious entity's success in identifying the actual location of a service entity is considered as a location privacy threat. Due to the open nature of the ubiquitous environment, a malicious entity may backtrack the path, through which service packets are traversed from consumers towards the service provider or vice versa. A malicious entity can also reach the actual location of the target legitimate one by following the flow direction of the service packet. The location of a service entity may be identified by stealing the network address and privacy-related other sensitive information about the location from the service packet. Moreover, sometimes a malicious entity observes the past location visits continuously and concludes the target entity.

In Fig. 2, we have shown a network segment having malicious entity M. It is divulging the location privacy of target CN-1 and SP-1, based on network address and movement history from the service packets. M maintains "Address Log" and "Movement Log" for CN-1 and SP-1. From these Logs, M able to make the service entity's "Location Profile" that leads M to reach the actual location of a legitimate service entity to identify and misuse their privacy.

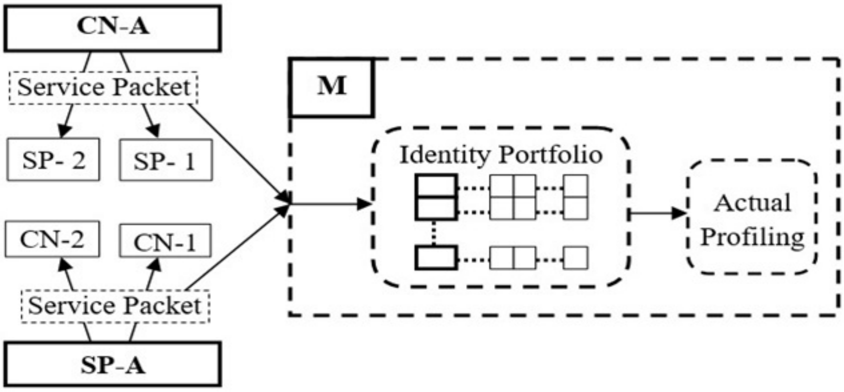

FIGURE 1. Identity privacy forging.

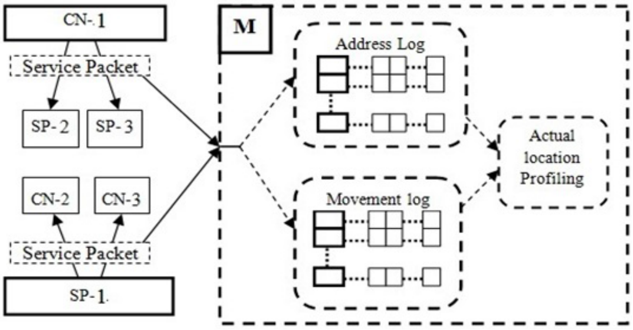

FIGURE 2. Location privacy forging.

C. BEHAVIOURAL PRIVACY THREAT

A malicious entity may identify the target SE by following its behavioural pattern with the continuous study of activity history and communication interests that are embedded in service packets. With one service entity's activity history, malicious entity maintains "Activity Log" with activity type (AT), classified activity description (CAD), activity association (AA) and activity time (At) with respect to parametric values in "Activity Log" for an individual service entity's identity as follows:

$$ActivityLog \Rightarrow \{Identity: \; A_1\{AT_1, CAD_1, AA_1, At_1\}\} \quad for, \; Activity \; A_1 \tag{3}$$

On the other hand, an adversary maintains the "Interest Log" by following the interest in service packets. For an individual SE's identity, the malicious entity maintains service type (ST), classified service description (CSD), service time (St) with parametric values in "Interest Log" as follows:

$$InterestLog \Rightarrow \{Identity: \; SPck_1\{ST_1, CSD_1, St_1\}\} \quad for, \; ServicePacket \; SPck_1 \tag{4}$$

In Fig. 3, we have shown a network segment having malicious entity M. It is breaking the behavioural privacy of CN-1 and SP-1. Here, M maintains "Activity Log" and "Interest Log" for CN-1 and SP-1 based on their activity history and interests in its service packets. Using these logs, M constructs the "Behavioural Profiling" for CN-1 and SP-1. From this "Behavioural Profiling", an adversary can either directly identify the target CN-1 and SP-1 or establishes a false attractive service communication with them. Through this false attractive service communication, M pretends that it needs some personal information about them to continue the communication. When attracted CN-1 or SP-1 gives that personal information, M constructs "User Profile" and reach to them.

V. NETWORK MODEL, OBJECTIVE, AND THREE-FOLD PRIVACY MODEL

In this section, we have described the network model and objective of our privacy preservation approach in the underlying network.

A. NETWORK MODEL

Our considered network is ubiquitous, where entities communicate with each other through an open platform in a dynamic way. Here consumer gets service from service provider unobtrusively anytime from anywhere. Service communications are done through an on-demand basis between consumer and service providers. A service provider gives services based on the requirement of the consumers. Every service entity in the network dynamically creates a community of trusted entities around itself with multi-hop trust relation, and it is called a trust circle. Using such a trust circle, one service entity introduces 2-degree anonymity to preserve three-fold privacy. In our proposal, the considered network is assumed to have control over the repetition of communicating packets or messages. Every network entity is free to join and leave the ubiquitous network at any time to get services from anywhere without leaving any footprints of their identity for possible misuses.

B. OBJECTIVE

Preservation of three-fold privacy during packet transfer between the communicating entities is a challenge in

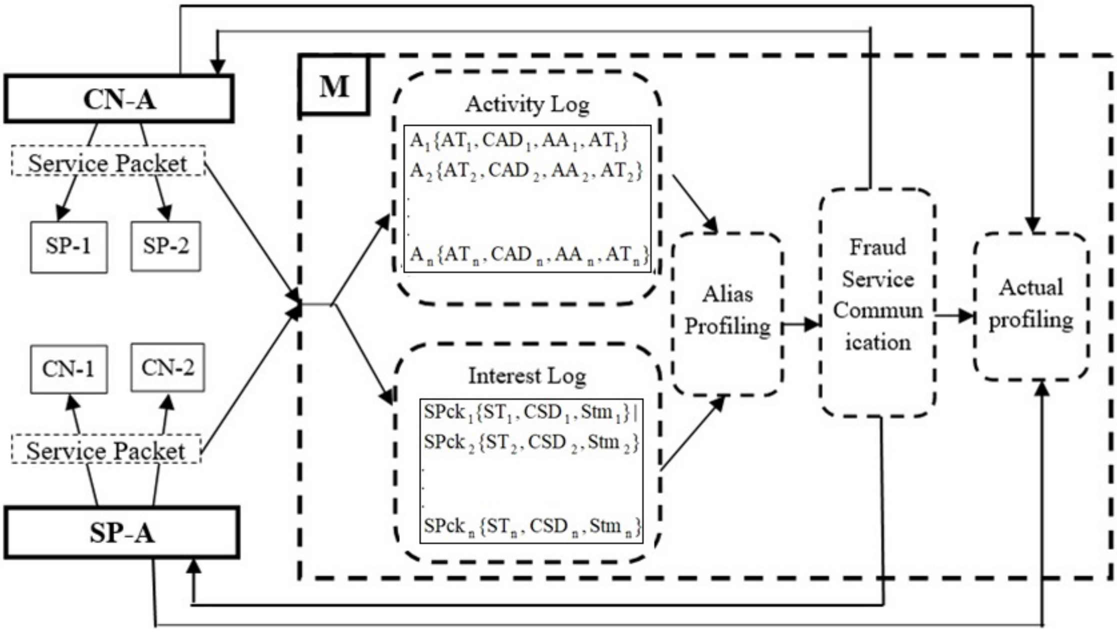

FIGURE 3. Behavioural privacy forging.

ubiquitous networks. Our privacy-preserving approach has two main objectives. The first one is to establish a trust circle around the service entity based on the context-dependent direct-indirect interaction and trust evaluation. The second one is the preservation of three-fold privacy with the two levels of anonymization by introducing 2-degree anonymity. 1st degree anonymity: "(p, β) sensitive k-anonymity with AlO" and 2nd degree anonymity: "anonymity by n-deviation with dummy messages between AcO and AlO" are introduced with the active participation of trust circle's members to provide privacy through the trusted community.

1) *Trust circle:* Our primary objective of trust circle establishment around every service communicable network user is to preserve their privacy by providing anonymity through their trust circle. A service entity selects different groups of trusted members from its trust circle for different service communications to provide anonymity. A large number of trusted entities in the trust circle of one facilitates more options for selecting different trusted groups for different service communications. Consequently, our focus is on expanding the trust circle from single-hop to multi-hop. We have focused on the hop-by-hop progressive trust model so that trust or belief level does not decay with the increase of hop-count with the non-interactive recommended entities. During the establishment and expansion of one service entity's trust circle, our objective is to preserve the identity privacy of entities by attribute matching in place of revealing actual identity.
2) *Three-fold privacy preservation with anonymity:* In a three-fold privacy preservation scheme, our main objective is to preserve privacy in terms of identity, location, and behaviour of service entity. The objective of our 1st degree anonymity is to protect AcO's identity, location address through (p, β) sensitive k-anonymity, and AlO is introduced to divert from behavioural tracking of AcO. On the other hand, the main focus of 2nd degree anonymity is to deviate malicious entity from the actual path to reach AcO for protecting the actual location, identity, and its behavioural tracking. In providing privacy with trust and anonymity, we have focused on the personalization between anonymity and trust for selecting AlO.

C. THREE-FOLD PRIVACY MODEL

Our privacy model contains the following state variables.

a) *Subject (Sub):* Subjects are the active stakeholders of the network.
b) *Object (O):* Trust Value. TC_{TP} record, Sensitive attributes, Non-sensitive attributes.
c) *Sensitive Object with different context (SO_C):* Sensitive attributes.
d) *Context (C):* Every sensitive attribute has a dependency on a context.
e) *Task (T):* A subject is allowed to access an object by performing a task. If a subject is not currently performing any task currently, It will be defined by the value Nil. A function CT: Sub -> T ∪Nil is defined, where CT(Subi) is the current task of subject Subi. Where tasks are :Trust evaluation, Malicious entity identification, TC_{TP} formation, (p, β) sensitive k-anonymous group, Alo selection, n deviation, Recommendation.
f) *Purpose (P):* Every task has to serve a certain purpose where, T-Purpose: T -> P

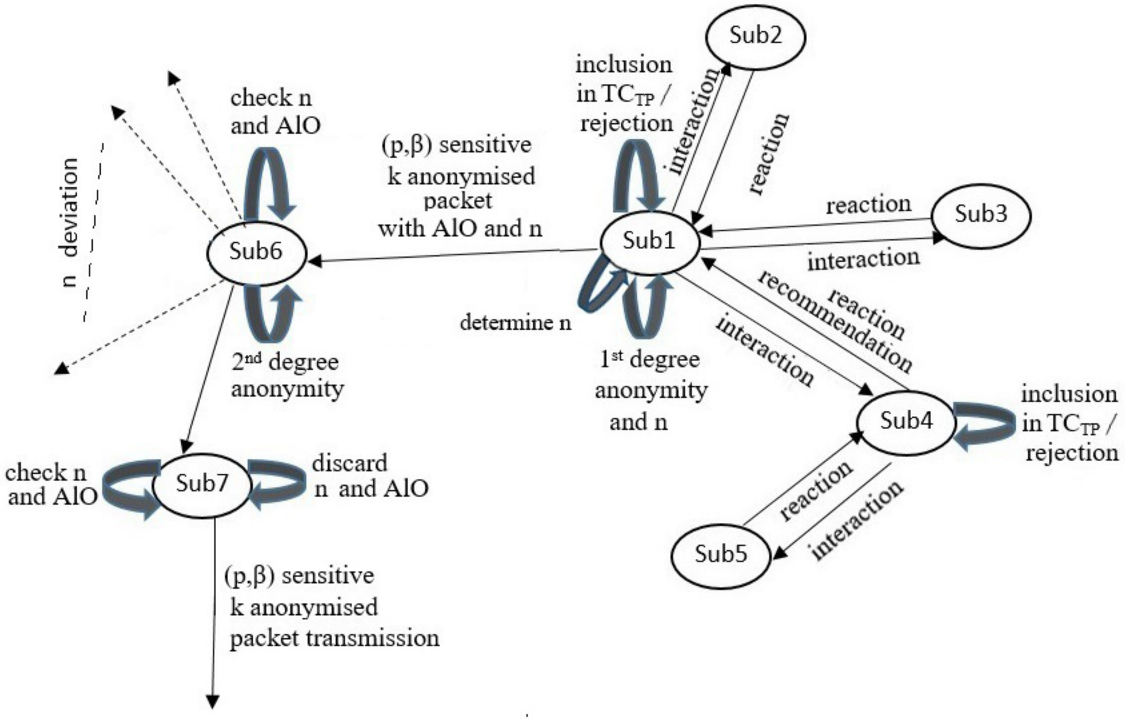

FIGURE 4. State model of privacy.

g) *Subject Role (Sub_R):* A Subject is categorised by different role for different purpose. A subject performs a task with a particular purpose. Sub_R: T -> P. Eg. Network user, user as AcO, user as AlO, user as intermediate nodes between AlO and AcO.

A privacy model for a network part with a particular source subject (Sub1) is depicted in Fig. 4. However. Sub1 is one of the representative members of the system. However, every subject may play different roles at the same time for different purposes. To better understand the privacy scenario, we have shown a privacy model for every single subject in Fig.5. In Fig. 4, Sub1 interacts with other subjects: Sub3, Sub4 including malicious one Sub2, who are the direct trustee of Sub1. Sub1 evaluates the trust of DTEs to accepts in the TC_{TP} or rejects based on the evaluated trust value. On the other hand, Sub4 recommends its trusted Sub5 for considering as RTE, and to evaluate for inclusion in Sub1's TC_{TP}. Sub1 either accepts or rejects Sub5 as RTE based on progressive trust property. Sub1 evaluates Sub5's trust as indirect trust.

We have shown all the privacy activities of a single subject/entity in Fig. 5. A subject forms its TC_{TP} of trusted members from interactions and recommendations of the direct and recommended trustees. An entity receives reactions of interactions for trust evaluation in Trust block and evaluates the trust. Based on computed trust, a trustee may be included in evaluating entity's TC_{TP} with trust >0.5 otherwise rejected as a malicious one. TC_{TP} with accepted trusted members is stored in the repository. When an entity generates a packet as source (AcO), it reports addresses of (p, β) sensitive k-anonymized group as source for its packet, instead of incorporating its own address as 1st degree anonymity. This anonymized group is selected from AcO's TC_{TP}. AcO selects AlO from (p, β) sensitive k anonymized group and sends the originated packet to AlO On the other hand, AcO determines the 'n' and appends it to the packet for 2nd degree anonymity. When an entity receives a packet, it checks for n and AlO. If it is an intermediate node of AlO and AcO, it will find 'n' and AlO and will process 2nd degree anonymity with n deviation. If a packet receiving entity found itself as an AlO, it discards the AlO and 'n' from the packet and transmits packet as AcO. If there is no AlO and 'n' in the received packet, then the packet receiving node id is an intermediate node between AlOs of source and destination of a communication. In this case, that receiving node just processes the packet and retransmit as per routing protocol.

VI. TRUST CIRCLE ESTABLISHMENT

In ubiquitous networks, service entities maintain their trust circle (TC_{TP}) depending on context-based direct and indirect interaction for a particular period, TP. A service entity (SE) reports a group of trusted entities from its TC_{TP} as the source. SE is considered as the trustor (TR) when it evaluates other entity's (direct or recommended trustee) trust to establish or to expand its TC_{TP}. SE considers a direct trustee (DTE) as a member of TC_{TP} based on direct trust evaluation by context-based interactions and recommendations. On the other hand, SE considers a recommended trustee (RTE) as

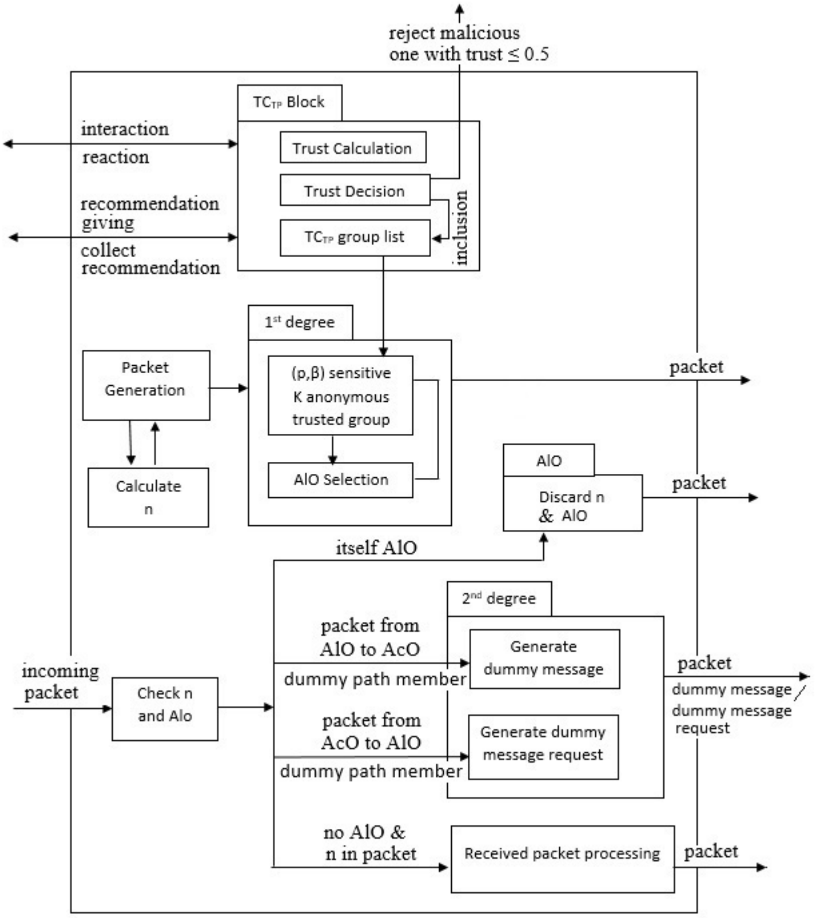

FIGURE 5. Privacy model.

a member of TC_{TP} based on indirect trust computation. All interactions of SE with DTE or RTE, and recommendation giving processes are done through attribute matching in place of revealing actual ID.

A. SELECTION OF DTE AS A TRUSTED MEMBER OF SE's TC_{TP}

SE evaluates the trust level for DTE based on direct and indirect trust computations. To evaluate the direct trust for DTE, SE measures the DTE's probability of being legitimate based on context-dependent positive and negative direct interactions. On the other hand, indirect trust is evaluated based on collaborative recommendations. The direct and indirect trust evaluation processes are described in this section.

For trust evaluation, positive interaction by an entity denotes the legitimacy of an entity, and negative interaction denotes the malicious behaviour. One DTE's probability of being legitimacy varies with the increase or decrease of the context-based positive interaction numbers. According to Bernoulli distribution, the probability of positive interaction is described as per "(5)". A discrete distribution that has two possible outcomes, n = 0 (failure), and n = 1 (success) with probability q = (1 − p), is described as Bernoulli's distribution [36].

$$\begin{aligned} &P_{leg}(Intr_q^c = true) = P; \\ &P_{leg}(Intr_q^c = false) = P - 1; \\ &for,\ q = 1, 2, \ldots tn \end{aligned} \tag{5}$$

$P_{leg}(Intr_q^c = true)$ is the measure of q^{th} interaction being positive with respect to legitimacy, where p is its probability. Every $Intr_q^c$ is independent of each other, and they have the same probability density function. Based on a context, positive interaction returns a true value, and negative interaction returns a false value. The binomial distribution provides a discrete probability distribution. The probability of obtaining 'pn' number of positive interactions that are having the true value of an entity with trust evaluating SE is $P_{leg}(pn = true|p)$. $P_{leg}(pn = true|p)$ follows the binomial distribution shown in "(6)".

$$P_{leg}(pn = true|p) = \binom{a}{b} p^{pn}(1-p)^{tn-pn=nm} \tag{6}$$

$P_{leg}(pn = true|p) \in [0.0,\ 1.0]$. SE's direct trust on DTE is directly proportional to $P_{leg}(pn = true|p)$ (see "(7)") .

$$[{}_{SE}(DT)_{DTE}]^t \propto P_{leg}(pn = true|p) \tag{7}$$

Comparing "(7)" and "(1)", $[{}_{SE}(DT)_{DTE}]^t$ is being decided by SE. On the other hand, SE computes the indirect trust ($[{}_{SE}(IT)_{DTE}]^t$) for DTE as per "(8)", using the recommendation ($[{}_{Ri}(RT)_{DTE}]^{tp_{comp}}$) of i^{th} recommender R_i for DTE and the acceptance rate $(Rate_{acc})_{DTE}$.

$$[{}_{SE}(IT)_{DTE}]^t = \sum_{i=1}^{n} [{}_{Ri}(RT)_{DTE}]^{tp_{cmp}} \times (Rate_{acc})_{DTE} \qquad for,\ tp_{cmp} > 0 \tag{8}$$

A trust value decreases with respect to time increase. R_i compares tp_{cur} with tp_{cmp}, at the time of forwarding $[{}_{Ri}(RT)_{DTE}]^{tp_{cmp}}$ for DTE to SE. If, $tp_{cur} = tp_{cmp}$, then R_i sends exactly the same value of its own computed trust for DTE as recommendation. On the other hand, if R_i finds that the tp_{cmp} is older than tp_{cur}, R_i sends decayed value of its own computed trust ($[{}_{Ri}(T)_{DTE}]^{tp_{comp}}$) for DTE as per "(9)". where, $e^{\alpha}(tp_{cur} - tp_{cmp})$ is the decay function and α is the decay rate ($\alpha > 0$).

$$[{}_{Ri}(RT)_{DTE}]^{tp_{cmp}} = \left\{ \begin{array}{l} [{}_{Ri}(T)_{DTE}]^{tp_{cmp}} \\ For, tp_{cmp} = tp_{cur} \\ \\ [{}_{Ri}(T)_{DTE}]^{tp_{cmp}} \times e^{\alpha}(tp_{cur} - tp_{cmp}) \\ For, tp_{cmp} < tp_{cur} \end{array} \right\} \tag{9}$$

The $(Rate_{acc})_{DTE}$ is evaluated by SE with its trust on R_i ($[{}_{SE}(T)_{Ri}]^{tp_{comp}}$) as per "(10)" with the consideration of tp_{cur} and tp_{comp}.

$$(Rate_{acc})_{DTE} = \left\{ \begin{array}{l} [{}_{SE}(T)_{Ri}]^{tp_{cmp}} \\ For,\ tp_{cmp} = tp_{cur} \\ [{}_{SE}(T)_{Ri}]^{tp_{cmp}} \times e^{\alpha}(tp_{cur} - tp_{cmp}) \\ For,\ tp_{cmp} < tp_{cur} \end{array} \right\} \tag{10}$$

SE computes the final trust ($[{}_{SE}(FT)_{DTE}]^{tp_{comp}}$) for DTE as per "(11)" using weight W. Here, W depends on application with limit $0.0 \leq W \leq 1.0$.

$$[{}_{SE}(FT)_{DTE}]^{tp_{comp}} = \left\{ \begin{array}{l} W \times [{}_{SE}(DT)_{DTE}]^{tp_{comp}} \ + \\ (1 - W) \times [{}_{SE}(IT)_{DTE}]^{tp_{comp}} \end{array} \right\} \tag{11}$$

Only those DTEs are considered as trusted member of TC_{TP} for which the value of $[{}_{SE}(FT)_{DTE}]^{tp_{cmp}} > 0.5$ (0.5=uncertainty point). The condition "TE==DTE" of Algorithm 1 describes the process of selecting DTE SE's TC_{TP}.

B. SELECTION OF RTE AS A TRUSTED MEMBER OF SE's TC_{TP}

One SE's trust circle (TC_{TP}) is extended from single-hop to multi-hop by considering RTE in TC_{TP} through our following evaluations.

Algorithm 1 Selection of DTE and RTE by SE in Its TC_{TP}

```
for each TE do
    if TE = = DTE then
        SE evaluate $P_{leg}(pn = true|p)$ for DTE based on context dependent interaction;
        SE measures ($[{}_{SE}(DT)_{DTE}]^{tp_{comp}}$) by $f\{\ P_{leg}(pn = true|p)\ \}$;
        SE computes ($[{}_{SE}(IT)_{DTE}]^{tp_{comp}}$) by $f\{\ [{}_{Ri}(RT)_{DTE}]^{tp_{comp}}\ ,\ (Rate_{acc})_{DTE}\ \}$;
        SE computes ($[{}_{SE}(FT)_{DTE}]^{tp_{comp}}$) by $f\{\ [{}_{SE}(DT)_{DTE}]^{tp_{comp}}\ ,\ [{}_{SE}(IT)_{DTE}]^{tp_{comp}}\ \}$;
        if $[{}_{SE}(FT)_{DTE}]^{tp_{comp}}$ > uncertain trust then
            $\text{TC}_{TP} \xleftarrow{member\ of} DTE$;
        end if
    end if
    if TE = = RTE then
        $\text{R}_{tr} \xrightarrow{recommendation\ for\ RTE} SE$;
        if $\left[{}_{R_{tr}}(RT)_{RTE}\right]^{tp_{comp}} \geq \left[{}_{SE}(T)_{R_{tr}}\right]^{tp_{comp}}$ then
            SE computes ($[{}_{SE}(FT)_{RTE}]^{tp_{comp}}$) by $f\{\ [{}_{SE}(RT)_{RTE}]^{tp_{comp}}\ ,\ Path_{weight}\ \}$;
        end if
        if $[{}_{SE}(FT)_{RTE}]^{tp_{comp}}$ > uncertain trust then
            $\text{TC}_{TP} \xleftarrow{member\ of} RTE$;
        end if
    end if
end for
```

SE's trusted entity in TC_{TP} recommends its own trusted entity (who are not in direct interaction with DTE and are not in TC_{TP} of SE) to SE. Recommendation giving trusted entities of SE in TC_{TP}, are considered as trusted recommenders (R_{tr}) only when the progressive trust acceptance condition is true. Since multi-hop entities are not in direct interaction, belief level may decay with the increase of hop-count. With this assumption, we have introduced progressive trust, following which R_{tr} forwarded $[{}_{R_{tr}}(RT)_{RTE}]^{tp_{comp}}$ is considered by SE, only when $[{}_{R_{tr}}(RT)_{RTE}]^{tp_{comp}} \geq [{}_{SE}T_{R_{tr}}]^{tp_{comp}}$. Here, $[{}_{R_{tr}}(RT)_{RTE}]^{tp_{comp}}$ is the recommendation trust for RTE by recommender R_{tr} and $[{}_{SE}T_{R_{tr}}]^{tp_{comp}}$ is the trust of SE for R_{tr}.

On receiving the R_{tr} forwarded recommendation, SE computes the final trust $[{}_{SE}FT_{RTE}]^{tp_{comp}}$ for RTE as per "(12)". See condition "TE==RTE" of Algorithm 1.

$$[{}_{SE}(FT)_{RTE}]^{tp_{comp}} = \left[{}_{R_{tr}}(RT)_{RTE}\right]^{tp_{comp}} + Path_{weight} \qquad For, \left[{}_{R_{tr}}(RT)_{RTE}\right]^{tp_{comp}} \geq \left[{}_{SE}(T)_{R_{tr}}\right]^{tp_{comp}} \tag{12}$$

$Path_{weight}$ is calculated by SE using $\left[{}_{SE}(T)_{R_{tr}}\right]^{tp_{comp}}$ as per "(13)". When $tp_{cur} = tp_{cmp}$, the value of $e^{\alpha}(tp_{cur} - tp_{cmp})$ becomes 1.

$$Path_{weight} = \left[{}_{SE}(T)_{R_{tr}}\right]^{t} p_{comp} \times e^{\alpha}(tp_{cur} - tp_{cmp}) \tag{13}$$

In "(12)", $[{}_{R_{tr}}(RT)_{RTE}]^{tp_{comp}}$ is exactly the same value of R_{tr}'s own computed trust for RTE when tp_{cur} =

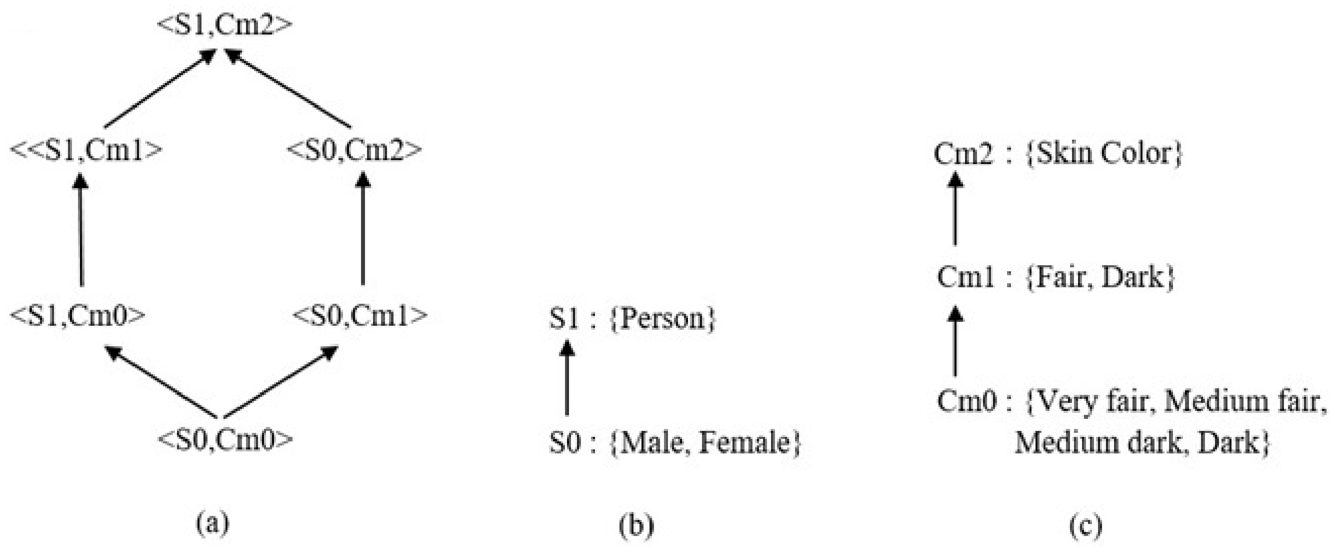

FIGURE 6. Generalization. In Fig. 6(a), generalization lattice of sex (S) and complexion (Cm) are shown. Fig. 6(b) and Fig. 6(c) is the domain generalization hierarchies for sex (S) and complexion (Cm), respectively.

tp_{cmp}. Otherwise it is decayed by time decaying function $e^{\alpha}(tp_{cur} - tp_{cmp})$ as per "(14)".

$$(Rate_{acc})_{DTE} = \left\{ \begin{array}{l} \left[R_{tr}(T)_{RTE}\right]^{tp_{cmp}} \\ For,\ tp_{cmp} = tp_{cur} \\ \left[R_{tr}(T)_{RTE}\right]^{tp_{cmp}} \times e^{\alpha}(tp_{cur} - tp_{cmp}) \\ For,\ tp_{cmp} < tp_{cur} \end{array} \right\} \quad (14)$$

If $[_{SE}(FT)_{RTE}]^{tp_{comp}}$ is greater than the uncertain trust value 0.5, that RTE become a member of SE's TC_{TP}.The condition "TE==RTE" of Algorithm 1 describes the process of selecting RTE SE's TC_{TP}.

VII. THREE-FOLD PRIVACY PRESERVATION WITH 2-DEGREE ANONYMITY

SE preserves its identity privacy, behavioural privacy, and location privacy at its own side by introducing 2-degree anonymity with the help of trusted entities in its established TC_{TP}. 1st degree anonymity is introduced by (p, β) sensitive k-anonymity with alias originator (AlO) and 2nd degree anonymity is introduced by n-deviation with dummy messages between AlO and AcO. Personalization is introduced in AlO selection between anonymity and trust.

2-Degree anonymity implies the two levels of anonymization. In 1st degree anonymity, the 1st level of anonymization is introduced through "(p, β) sensitive k-anonymity with AlO". It preserves the threefold privacy in the context of identity, location, and behaviour. But in the scenario where adversary continuously studies the communications AlO and AcO, behavioural privacy may be compromised, and that can affect location privacy also. That is why a 2nd level of anonymization is introduced in-between AcO and AlO. In 2nd degree anonymity, the 2nd level of anonymization is introduced through n-deviation with dummy messages.

A. 1st DEGREE ANONYMITY: (p, β) SENSITIVE k-ANONYMITY WITH AlO

In this scheme, SE introduces anonymity by (p, β) sensitive k-anonymity and an alias originator (AlO) for preserving privacy in service communication with the help of its TC_{TP}. When SE needs to send a service packet of j^{th} service communication within a period TP, it selects its j^{th} "(p, β) sensitive k-anonymity" group and its AlO from TC_{TP}. Here, k number of (p, β) sensitive trusted entities, that are having a group of non-sensitive attributes with identical values are reported as packet source in place of a single identity of the actual originator (AcO). These k trusted entities have a minimum p number of distinct sensitive attribute values among at least β number of distinct attribute types that are relevant to a context. SE considers one trusted entity from this "(p, β) sensitive k-anonymity" satisfied group to act as AlO. In our approach, anonymization is a process that transforms a table of records to its conforming "(p, β) sensitive k-anonymity" model. This process is employed in the generalization of quasi-identifier attributes' values. The generalization of a quasi-identifier attribute replaces the actual value of the attribute with less specific but semantically consistent with the original value. Samarati [21] introduced generalization lattice, based on the generalization hierarchy of two or more non-SA from the QI group. The highest level of generalization is called lattice height. In Fig. 6(a), the height is 3 for generalization lattice of sex (S) and complexion (Cm). The intermediate levels of generalization hierarchy are denoted as intermediate generalization heights. Fig. 6(b) and Fig. 6(c) is the domain generalization hierarchies for sex (S) and complexion (Cm), respectively, and the domains are S0, S1 for sex and Cm0, Cm1, Cm2 for the complexion.

Some anonymity related definitions are as follows:

- *Sensitive Attributes (SA):*
  Attributes that are unspecified to a malicious entity and whose values (SA-values) need to be confined to preserve privacy.
- *Quasi Identifier (QI):*
  Let Tb$(A_1, \ldots, A_m)$ is a table of entities' records $\{Rcd_1, Rcd_2, \ldots, Rcd_q\}$, having a set of non-sensitive attributes $\{NA_1, NA_2 \ldots, NA_r\}$, where $\{NA_1, NA_2 \ldots, NA_r\} \in \{A_1, \ldots, A_m\}$. That $\{NA_1, NA_2 \ldots, NA_r\}$ set is considered as quasi identifier when it re-identifies entities by linking external information. A QI-group is the set of records having identical values for QI attributes in table Tb.
- *k-anonymity:*
  Let Tb$(A_1, \ldots, A_m)$ is a table of entities' records $\{Rcd_1, Rcd_2, \ldots, Rcd_q\}$ and QI is the quasi identifier with non-sensitive attributes $\{NA_1, NA_2 \ldots, NA_r\}$ of Tb, where $\{NA_1, NA_2 \ldots, NA_r\} \in \{A_1, \ldots, A_m\}$. Tb is considered as k-anonymity satisfied table, iff each QI group has at least k number of records with identical values of QI attributes.
- *(p, β) sensitive k-anonymity:*
  Tb$(A_1, \ldots, A_m)$ is a table of entities' records and QI is the quasi identifier of it with non-sensitive attributes $\{NA_1, NA_2 \ldots, NA_r\}$, where $\{NA_1, NA_2 \ldots, NA_r\} \in \{A_1, \ldots, A_m\}$. Tb is considered as "(p, β) sensitive k-anonymity" satisfied table, iff it satisfies k-anonymity and at least p number of different SA-values among β number of distinct SA-types that are relevant to a context in each QI-group of Tb.

For example, Table 3 is a (p = 2, β = 2) sensitive k = 3 anonymity satisfied table where attribute "Zip Code" is already generalized up to one height (last digit). Each QI group deals with three records that are having similar non-SA-values and satisfies k = 3 condition. Each QI group has two different SA types (Medical Disease and Blood Group) under a relevant context. Consequently, the condition β = 2 is satisfied. Among β = 2 number of SA-types, each QI group has two different SA-values that satisfy p = 2 condition. It is hard to identify Bob for malicious Alice with two different SA-types, which prevent the link between two SA-values. If Alice guesses that Bob is a patient, even then, Alice could not identify whether Bob is suffering from any medical disease as a patient or not with the mentioned blood group.

TABLE 3. (p = 2, β = 2) sensitive k = 3 anonymity satisfied table.

Touple No.	Age	Country	Zip Code	SA Type	SA Value
1	1-25	America	1427*	Medical Disease	Viral Infection
2	1-25	America	1427*	Blood Group	A+
3	1-25	America	1427*	Blood Group	A+
4	51-75	America	1427*	Medical Disease	Flue
5	51-75	America	1427*	Blood Group	AB-
6	51-75	America	1427*	Blood Group	O+
7	26-50	America	1427*	Medical Disease	Heart Disease
8	26-50	America	1427*	Medical Disease	Heart Disease
9	26-50	America	1427*	Blood Group	AB+
10	76-100	America	1427*	Medical Disease	Flue
11	76-100	America	1427*	Medical Disease	Cancer
12	76-100	America	1427*	Blood Group	O-

In our work, SA-values deal with the information in Table 4 that describes the category, sensitivity, and weight.

TABLE 4. Sensitive attribute value.

SA-Type (β)	SA-Value	Category (P+)	Sensitivity	Weight (α)
Medical Disease	Cancer, Organ failure.	A	Top Secret	0
	Heart Disease, Malaria.	B	Secret	1/3
	Flue, Asthma.	C	Medium Secret	2/3
	Viral Infection, Acidity.	D	Less Secret	1
Blood Group	AB-, A-	A	Top Secret	0
	O- , B-	B	Secret	1/3
	AB+ ,r A+	C	Medium Secret	2/3
	O+, B+	D	Less Secret	1
Bank Account	Diamond, Platinum	A	Top Secret	0
	Gold, Silver	B	Secret	1/3
	Bronze , Iron	C	Medium Secret	2/3
	Green, Basic	D	Less Secret	1

- *Category (p+):*
  Partition among SA-values with different sensitivity.
- *Weight (α):*
  Weight to categorical SA-value is decided as per "(15)" following [34]. Sensitive attributes are denoted by S(SA) = $\{SA_1, SA_2 \ldots, SA_s\}$, where sensitivity decrease with increase of i ($1 \leq i \leq s$).

$$Weight(SA_i) = \frac{i-1}{s-1} \quad For, \ 1 \leq i \leq s$$
$$Weight(SA_s) = 1 \tag{15}$$

In the course of generalization to satisfy anonymity conditions, distortion in data is introduced. But our objective is to keep the distortion at the minimum level with desired level of anonymity.

- *Distortion:*
  If a value has been generalized to a more general value in generalization hierarchy, then there is a distortion of that attribute value. When the value of an attribute is not generalized, there is no distortion. Distortion increases with the increase in height of generalization, starting

TABLE 5. (p = 2, β = 2) Generalized table of Table3.

Touple No.	Age	Country	Zip Code	SA Type	SA Value
1	1-50	America	1427*	Medical Disease	Viral Infection
7	1-50	America	1427*	Medical Disease	Heart Disease
8	1-50	America	1427*	Medical Disease	Heart Disease
4	51-100	America	1427*	Medical Disease	Flue
10	51-100	America	1427*	Medical Disease	Flue
11	51-100	America	1427*	Medical Disease	Cancer
2	1-50	America	1427*	Blood Group	A+
3	1-50	America	1427*	Blood Group	A+
9	1-50	America	1427*	Blood Group	AB+
5	51-100	America	1427*	Blood Group	AB-
6	51-100	America	1427*	Blood Group	O+
12	51-100	America	1427*	Blood Group	O-

from 0 (when there is no generalization). If the value of a non-sensitive attribute is generalized one level up, the generalization height increased by 1.

Our "(p, β) sensitive k-anonymity" introduces less distortion in comparison to p, (p, α), p+, and (p+, α) sensitive k-anonymity. For example, in Table 3, considered satisfying conditions are p = 2, p+ = 2, α = 1, and β = 2. As per our "(p, β) sensitive k-anonymity" approach, Table 3 achieves the satisfying condition (p = 2, β = 2, and k = 3) only with this one height generalization of zip-code. But in case of p, (p, α), p+, and (p+, α) sensitive k-anonymity approaches, p, p+, and α conditions are evaluated under each SA-type separately. Consequently, Table 3 is not achieving a satisfying condition for them. To achieve the satisfying condition for p, (p, α), p+, and (p+, α) sensitive k-anonymity approaches, we need more generalization, which results in more distortion. If again, we generalize sensitive attribute "Age" of Table 3 to one height then we get Table 5, which achieves the satisfying condition for p, (p, α), p+, and (p+, α) sensitive k-anonymity.

If there are identical SA-values under every single SA-type, there is no distinct value to provide anonymity under each SA-type. Only one type (p = 1) of the SA-Values with a single category (p+ = 1) exists under each separate SA-type. It violates all the (p, α), p+, (p+, α) conditions as the value of p and p+ must be greater than 1, so that it can introduce at least two different sensitive values and categories. But our "(p, β) sensitive k-anonymity" approach supports identical SA-values under every single SA-type, as it supports combined consideration of β number of SA-types. Algorithm 2 describes "(p, β) sensitive k-anonymity" process.

Algorithm 2 (p, β) Sensitive k-Anonymity

```
Initialize $H_{cur}$ to 0
Set the value of p, k, and $\beta$
D ← each unique $S(NAv)_i$
while C $\neq$ NULL do
    for all $Rcd_i$ in C do
        $H\left[(NA_j)_{Rcdi}\right] \leftarrow H_{cur}$
    end for
    for all $S(NAv)_i$ in D do
        Count $O_{S(NAv)_i}$ by comparing every $Rcd_i$ having similar $S(NAv)_i$ in C
        Build $i^{th}$ QI group with similar $S(NAv)_i$
        Count $p_{(QI)}$ in a QI group of similar $S(NAv)_i$
        Count $\beta_{(QI)}$ in a QI group of similar $S(NAv)_i$
        if $p_{(QI)} \geq k$ , $p_{(QI)} \geq p$, and $\beta_{(QI)} \geq \beta$ then
            Remove record having $S(NAv)_i$ from C
            Remove $S(NAv)_i$ from D
        end if
    end for
    if D$\neq$ NULL then
        Select NAj to generalize one height for all $Rcd_i$ in C;
        Generalize $NA_j$ to its next height;
        $H_{cur} = H_{cur} + 1$;
    end if
end while
```

An actual service packet originator (AcO) selects j^{th} quasi identifier group $\left[QI_{(p,\beta).k}\right]_j$ from "(p, β) sensitive k-anonymity" satisfied TC_{TP} for j^{th} service communication. AcO selects one trusted member from that $\left[QI_{(p,\beta).k}\right]_j$ group as alias originator (AlO). AcO sends a service packet to AlO for broadcasting it as the real originator. All identities of trusted $\left[QI_{(p,\beta).k}\right]_j$ are reported as a source of service packet. On receiving that service packet, AlO finds its own identity among all appended identities and routes this service packet acting like AcO (see Algorithm 3).

1) PRIVACY PRESERVATION WITH 1st DEGREE ANONYMITY WITH AlO

The proposed (p, β) sensitive k-anonymity with AlO preserves different types of privacies in the following way:

- *Identity Privacy Preservation:*
  Since all identities of "(p, β) sensitive k-anonymity" group are reported as a source in service packet and in place of AcO, AlO routes the service packet; malicious entity can not divulge AcO's identity by interpreting the service packet's source. If one malicious entity tries to interpret the AcO's identity, it will get the identities of "(p, β) sensitive k-anonymity" group including AlO.
- *Location Privacy Preservation:*
  Since identities of "(p, β) sensitive k-anonymity" group from TC_{TP} is reported as source of the packet, one

Algorithm 3 Service Packet Routing by AIO

```
if $p_{(QI)} \geq k$ , $p_{(QI)} \geq p$, and $\beta_{(QI)} \geq \beta$ then
    for $j^{th}$ service communication is true do
        AcO select $\left[QI_{(p,\beta).k}\right]_j$ from $TC_{TP}$
        AcO appends all identities of $\left[QI_{(p,\beta).k}\right]_j$ to service packet as source
        AcO selects AlO from $\left[QI_{(p,\beta).k}\right]_j$
        AcO sends service packet to AlO
        if at AlO, service packet is true then
            AlO checks own identity in reported $\left[QI_{(p,\beta).k}\right]_j$ as source in service packet
            if AlO $\epsilon$ $\left[QI_{(p,\beta).k}\right]_j$ then
                AlO routes service packet acting like AcO
            end if
        end if
    end for
end if
```

malicious entity will be directed to the location of "(p, β) sensitive k-anonymity" group in trying to reach the location of AcO. If one malicious entity tries to locate the packet originator, the malicious entity reaches the AlO in place of AcO, as AlO is routed packets like actual originator.

- *Behavioural Privacy Preservation:*
  If one malicious entity tries to follow the behavioural pattern of AcO, it will follow the behaviour of the identities of [QI(p, β).k]j or AlO. As a consequence, a malicious entity can not break the behavioural pattern of AcO by interpreting the packet source. But with the continuous study of communication between Al and AcO, behavioural privacy may be compromised. This possibility necessitates the 2nd degree of anonymization.

B. 2nd DEGREE ANONYMITY: ANONYMITY BY n-DEVIATION WITH DUMMY MESSAGES

A malicious entity reaches to the AlO in place of AcO in our proposed "(p, β) sensitive k-anonymity with AlO" scheme. With the continuous study of behaviour and communication pattern between AlO and AcO, a malicious entity may get knowledge about AcO. With this knowledge, a malicious entity may reach to AcO by following the packet flow direction from AlO to AcO or by backtracking the packets flow from AcO to AlO. To prevent the malicious entity from reaching the AcO, we have incorporated "n-deviation with dummy message" between AcO and AlO to our "(p, β) sensitive k-anonymity with AlO" scheme.

In this approach, every intermediate trusted node (Im_i) initiates n-deviation with dummy messages. When intermediate Im_i forwards a service packet traversing from AlO to AcO, it sends dummy messages to its trusted neighbours, with the satisfying condition 'n'. Consequently, a malicious entity deviates to dummy branches following the direction of the service packet flow. On the other hand, when Im_i forwards service packet that traverses from AcO to AlO, it sends dummy message request to trusted neighbours for getting back dummy messages. As a consequence, a malicious entity is deviated from backtracking the actual traversed path of the service packet. Similarly, neighbours of Im_i continue this process until the traversed hop-count becomes equal to the hop-count between AlO and AcO. Dummy message traverse trough trusted neighbours in acyclic fashion. To decide 'n', AcO defines an adjustable ambiguity factor (A_{factor}) for j^{th} service communication. Depending on A_{factor}, AcO calculates the 'n' for a node as per "(16)".

$$n = \lfloor \frac{1}{1 - A_{factor}} \rfloor \quad For,\ 0.0 \leq A_{factor} \leq 1.0 \tag{16}$$

Ambiguity in finding the path is directly proportional to 'n'. Ambiguity depends on A_factor as per "(17)", Where, $A_{factor} \propto n$.

$$Ambiguity = \left\{ \begin{array}{l} Low\ Ambiguity \\ For,\ 0.0 < A_{factor} < 0.5 \\ \\ Medium\ Ambiguity \\ For,\ 0.5 \leq A_{factor} < 0.8 \\ \\ High\ Ambiguity \\ For,\ 0.8 \leq A_{factor} < 1.0 \\ \end{array} \right\} \tag{17}$$

After receiving of any service packet, Im_i compares the 'n' with its N_{nbr} to create dummy branches as follows:

$$\begin{aligned} Dummy\ branch &= n, \quad For,\ N_{nbr} \geq n \\ Dummy\ branch &= N_{nbr}, \quad For,\ N_{nbr} < n \end{aligned} \tag{18}$$

The higher value of 'n' introduces more anonymity with more deviation between AlO and AcO. In Fig. 7, 1st level of n-deviation in a network part is represented. Here we have considered a partial network consisting of AcO, AlO, and every intermediate node (Im_i node: 1, 2, 3, 4, 5) in-between AcO and AlO. Every Im_i creates n-deviation with dummy messages, where n = 3. The value of N_{nbr} for node1 is 4, for node2 is 2, for node3 is 3, for node4 is 2, and for node5 is 5. Consequently, dummy branches from node1 are 3, from node2 is 2, from node3 is 3, from node4 is 2, and from node5 is 3. When Im_i forwards a service packet that traverses from AlO to AcO, Im_i creates dummy branches by sending dummy messages to neighbours as shown in Fig. 7(a). On the other hand, when Im_i forwards service packet that traverses from AcO to AlO, Im_i requests neighbours so that they can send back dummy messages as in Fig. 7(b) (see Algorithm 4).

1) PRIVACY PRESERVATION IN 2nd DEGREE ANONYMITY

The proposed Anonymity by n-deviation with dummy messages preserves different types of privacies in the following way:

- *Identity Privacy Preservation:*
  Since malicious, the packets from them deviate from the actual path leading to AcO by dummy branches with

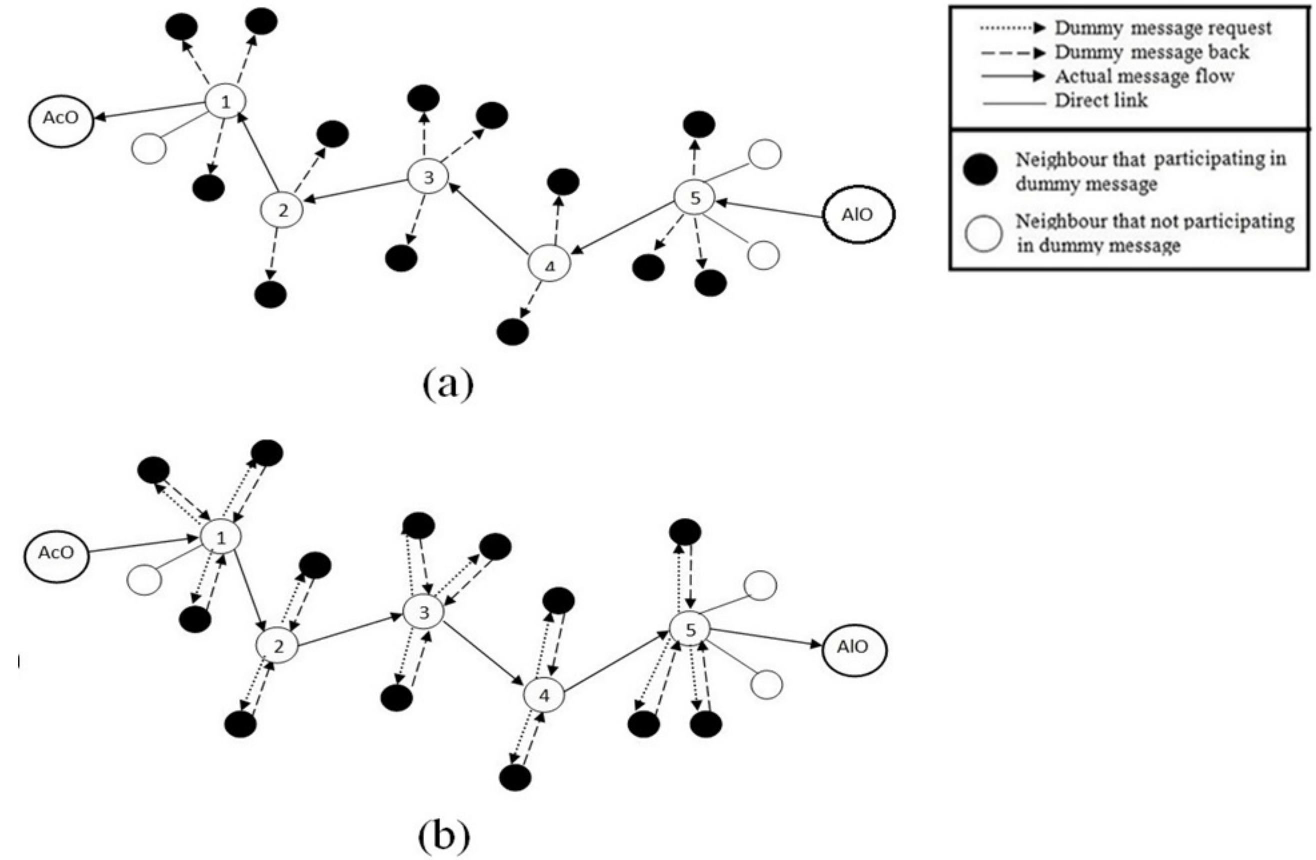

FIGURE 7. 1^{st} level of n-deviation in a network part which consists of AcO, AlO, and intermediate nodes (Imi node: 1, 2, 3, 4, 5). Every Imi creates n-deviation with dummy messages, where n = 3.

Algorithm 4 Anonymity With n-Deviation by Im_i

```
for $j^{th}$ service communication is true do
    Decides $A_{factor}$ based on ambiguity requirement
    n = f{ $A_{factor}$ } at $Im_i$
    if service packet is received at $Im_i$ then
        if service packet is traversing from AlO to AcO then
            dummy branch at $Im_i = f\{n, N_{nbr}\}$ by sending dummy message
        end if
        if service packet is traversing from AcO to AlO then
            Dummy branch at $Im_i = f\{n, N_{nbr}\}$ by sending dummy message request
        end if
        Forwards service packet to next-hop node in the actual message flow path between AcO and AlO
    end if
end for
```

Algorithm 5 Personalized AlO Selection

```
AlO selects $TC_{TP}$ in a time period TP
for $j^{th}$ service communication is true do
    AlO Select $W_I$
    for $K_i = 1; K_i \leq m; K_i++$ do
        $Pval = f\{W_I, T_{ki}, Hcount_{ki}\}$ for $Entity_{Ki}$
    end for
    if Pval of $Entity_{Ki}$ is highest then
        AlO = $T_{Ki}$
    end if
end for
```

dummy messages, and the malicious entity will not get the identity of AcO.

- *Location Privacy Preservation:*
  Since packets from malicious entities can not reach to the actual location of AcO due to path diversions by dummy messages, location privacy will be preserved.
- *Behavioural Privacy Preservation:*
  Since the packets from malicious entities are misguided by diversified dummy messages, malicious entities can not identify behaviour of AlO or AcO. Hence the behavioural privacy of AcO will be preserved.

C. PERSONALIZATION BETWEEN ANONYMITY AND TRUST IN AlO SELECTION

In our approach, the mechanism of building trust circle (TC_{TP}) is expanded with the multi-hop trust relationships between entities. From such multi-hop distant trusted members of ''(p, β) sensitive k-anonymity'' group, AcO selects AlO. Higher hop-count between AlO and AcO introduces higher anonymity to reach AcO from AlO. As hop-count increases between AlO and AcO, the deviation from intermediate nodes increases to reach AcO, following n-deviation. Consequently, anonymity increases with hop-count. On the other hand, an entity with a high trust value is very important for ''(p, β) sensitive k-anonymity with AlO'' in comparison to another entity with lower trust but higher hop-count. Keeping these two scenarios in mind, we have provided the personalization to choose the importance of 'trust' and the anonymity factor 'hop-count' in AlO selection. The personalized value of importance-weight (W_I) gives priority to the entity's trust or hop-count. AcO computes the priority value (Pval) of entities in its j^{th} quasi identifier group ($\left[QI_{(p,\beta).k}\right]_j$) with the help of trust, hop-count and W_I. The entity, whose Pval value is computed as highest among all entities in $\left[QI_{(p,\beta).k}\right]_j$ group is considered as AlO (see Algorithm 5). If trust is T_{Ki} and hop-count is $Hcount_{Ki}$ of Ki^{th} entity $Entity_{Ki}$ in $\left[QI_{(p,\beta).k}\right]_j$ group of AcO, then Pval is derived through personalized value of (W_I) as per ''(19)'' for j^{th} service communication.

$$Pval = T_{Ki} \times \left(\frac{1}{W_I}\right)^{Hcount_{Ki}}$$
$$For, \quad 0.5 \leq W_I \leq 1.0 \tag{19}$$

In our consideration, $Entity_{Ki}$ is considered in TC_{TP} with $0.5 < T_{Ki} \leq 1.0$ and importance weight (W_I) that varies from 0.5 to 1.0. For $W_I = 0.5$ (lower bound), total importance goes to one's $Hcount_{Ki}$, and T_{Ki} does not have any importance. Pval becomes higher for an entity having a higher $Hcount_{Ki}$ at $W_I = 0.5$. With the increase of W_I's value from 0.5 to 1.0, the importance of T_{Ki} increases. For $W_I = 1.0$ (higher bound), total importance goes to T_{Ki} of one and it's $Hcount_{Ki}$ does not have any importance. Pval becomes greater for an entity having greater trust at $W_I = 1.0$, no matter whether it has high hop-count or low.

Lemma 1: For $W_I = 1.0$, the value of Pval is always greater for a node with higher T_{Ki} than another node with lower T_{Ki} for any $Hcount_{Ki}$

Proof: If $W_I = 1.0$, then $(\frac{1}{W_I})^{Hcount_{Ki}} = 1.0$ and $T_{Ki} \times (\frac{1}{W_I})^{Hcount_{Ki}} = T_{Ki}$. As a result, Pval $= T_{Ki}$. In that case, with the higher value of T_{Ki}, Pval of an entity will be higher and vice versa. Consequently, Pval is always greater for an entity with higher T_{Ki}. ■

Lemma 2: For $W_I = 0.5$, the value of Pval is greater for a node with higher $Hcount_{Ki}$ than a node with lower $Hcount_{Ki}$ for any value of T_{Ki}.

Proof: If $W_I = 0.5$ and $Hcount_{Ki} = h$ then the value of Pval is $\left(T_{Ki} \times 2^h\right)$. For a particular value of T_{Ki}, $\left(T_{Ki} \times 2^h\right) < \left(T_{Ki} \times 2^{(h+1)}\right) \ldots < \left(T_{Ki} \times 2^{(h+2)}\right)$. It implies that Pval is greater for any greater value of $Hcount_{Ki}$ with T_{Ki}. Since, $0.5 < T_{Ki} \leq 1.0$, the relation is true for any value of T_{Ki}. ■

VIII. COMPLEXITY ANALYSIS

Computing direct trust of DTE is the order of $\binom{tn}{pn}$, since p is a probability value between 0 and 1. So the cost of direct trust computation is O(C(tn,pn)) which means O(n choose k). The complexity of computing indirect trust of DTE by SE is O(n_R), where n_R is the total number of recommendations. The complexity of computing the trust for RTE is O(1). (for one recommendation path). In [29], the optimal p sensitive k-anonymity problem is proved as an NP-hard problem. As a consequence, it is easy to say that optimal (p, β) sensitive k-anonymity is NP-hard considering the situation where p numbers of different sensitive values are from β numbers of different sensitive attribute types under a context of interactions. The complexity of AlO selection is O(k), where k is the number of considered QI group members. The communication cost of an intermediate node between AlO and AcO is O(n), where it needs to communicate with n neighbours to initiate n deviation, and needs to transmit the actual packet.The communication cost of an AcO is O(1), where it only needs to communicate with a next-hop neighbour. The communication cost of a recommender for a particular trustee is O(1). On the other hand, the communication cost of a trust evaluating trustor is O(n_R).

IX. SIMULATION

To evaluate the performance of building one entity's social trust circle, we have used the ''Facebook'' data set [1]. The performance analysis of our proposed trust circle (TC_{TP}) establishment algorithm is carried out concerning reachability, based on the collected ''Facebook'' data from two universities, Amherst and Colgate.

- *Reachability:* Reachability is defined as the actual number of trusted entities with respect to the total possible number of trusted entities of all members in the entire network.

Our TC_{TP} establishment scheme greatly increases the reachability with respect to hop-count as it is extended from single-hop to multi-hop based on recommendation.

In Fig. 8(a) and Fig. 8(b), the depicted reachability measures are showing the high performance of our proposed TC_{TP} establishment over the existing circle-establishing scheme. The reachability of the existing algorithm is consistently low with respect to the hop count increase from hop 1 to hop 5, as it does not support the multi-hop extension of the social circle. From hop 1 to hop 5, the reachability of our TC_{TP} establishment scheme is 4.00, 48.00, 98.00, 98.50, and 99.00, respectively. Multi-hop TC_{TP} of every entity in the network includes all most 'all possible trusted entities' at hop 3 expansion. On the other hand, the reachability of existing circle establishment is 4.00 at hop 1, hop 2, hop 3, hop 4 as well as hop 5.

In Fig. 9(a) and Fig. 9(b), we have compared the reachability of our proposed TC_{TP} establishment scheme with the reachability of ID-based circle establishment under the condition of multi-hop recommendation. Here, we can see that

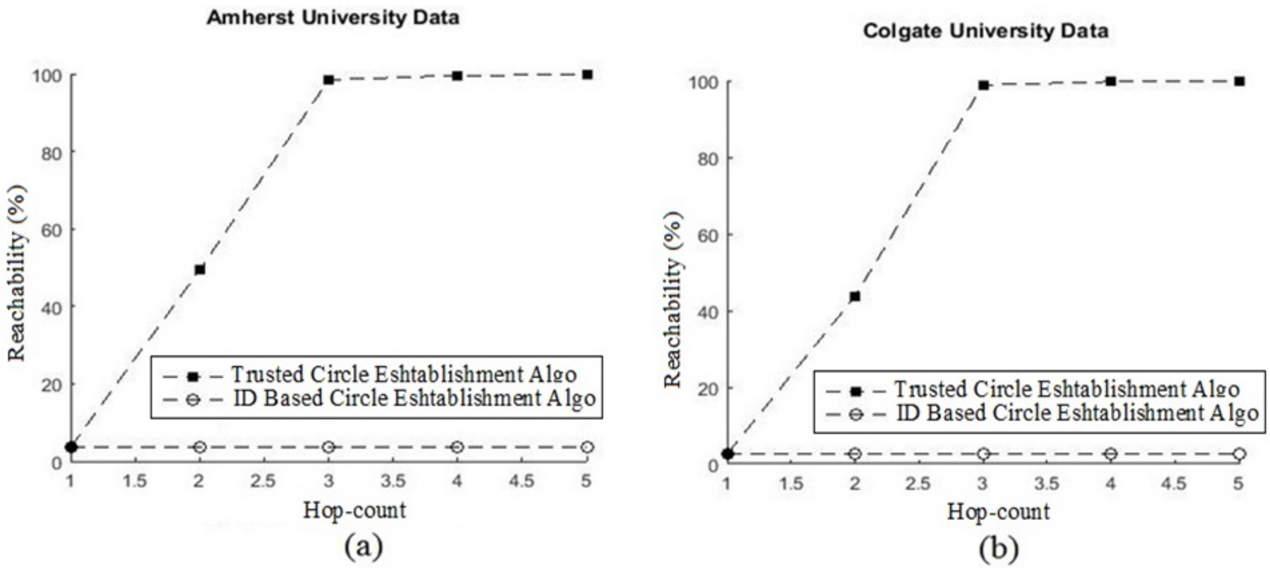

FIGURE 8. Reachability vs. hop-count. In Fig. 8(a) and Fig. 8(b), the depicted reachability measures are showing the high performance of our proposed trust circle establishment over the existing circle-establishing scheme.

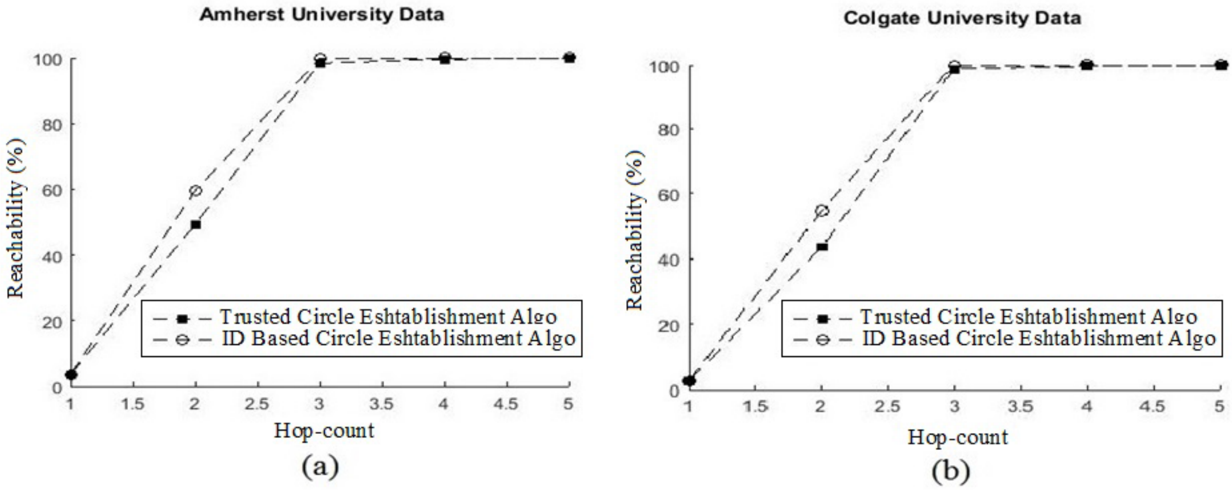

FIGURE 9. Reachability vs. hop-count. In Fig. 8(a) and Fig. 8(b), the depicted reachability measures are showing the high performance of our proposed trust circle establishment over the existing circle-establishing scheme under the condition of multi-hop recommendation.

the reachability of 'ID-based circle establishment' is a little bit higher than our proposed TC_{TP} establishment scheme with respect to hop count increase under the condition of multi-hop recommendation. The reachability measure of the ID-based scheme is 59.51%, 99.82%, 100% at hop 2, hop 3, and hop 4 respectively, whereas the reachability measure of our trust circle scheme is 49.32%, 98.42%, 99.61% which is a little bit lesser than ID-based scheme. The main reason for that is our scheme filters out "unqualified" direct or multi-hop entities based on trust measure. One trustor considers a recommended trustee (RTE) as a trusted member in TC_{TP} only when the trust level of RTE is greater than equal to the recommender trusted one. This progressive trust acceptance results in a qualified trust chain to multi-hop trusted entities. Otherwise, trust value can decrease in a multi-hop trust relationship.

On the other hand, our experiment for anonymity is based on a real "ADULT" database [31], which is publicly available at the UC Irvine Machine Learning Repository. The ADULT database contains 48842 tuples. We eliminated the unqualified tuple with unknown values and considered 45222 tuples. Each tuple describes personal information. Our simulation treats seven non-sensitive attributes: age, workplace, education, marital status, race, gender, and native-country of the "ADULT" database attributes. In Table 6, considered attributes are described with the type of each attribute, the number of distinct values for each attribute, and the height of the generalization hierarchy for each attribute. We added two-column, one of which describes types of sensitive attribute and describes sensitive attribute values. We consider three sensitive attributes: "Medical Disease", "Blood

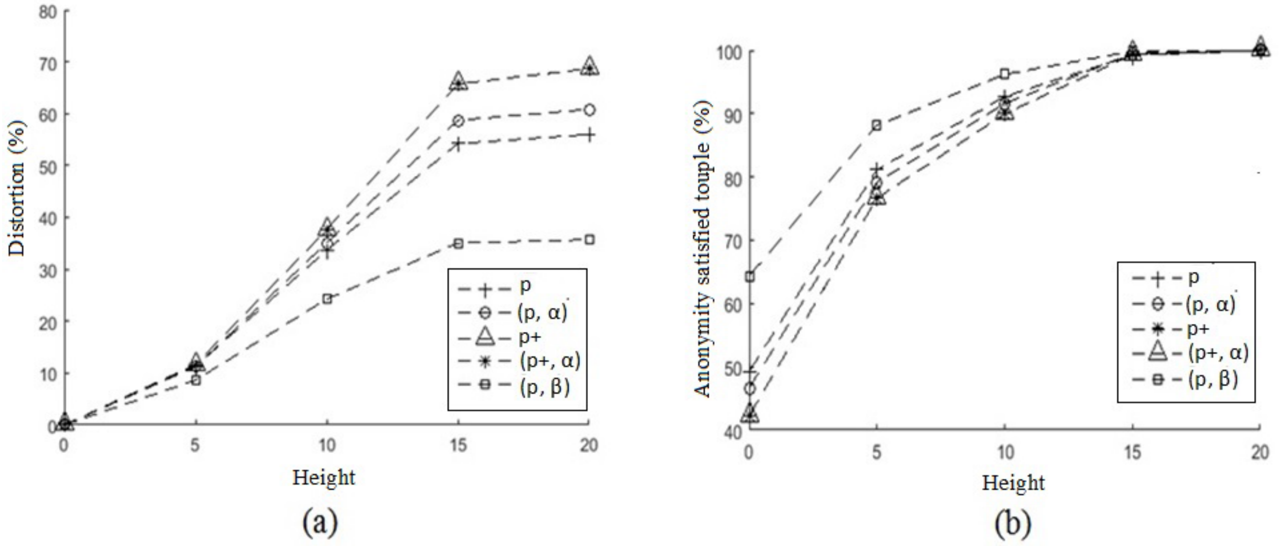

FIGURE 10. It shows the distortion ratio of our proposed (p, β) sensitive k-anonymity approach in comparison to existing anonymity approaches with respect to generalization height.

TABLE 6. Attribute description.

Attribute	Type	Distinct Value	Height
Age	Numeric	74	4
Workclass	Categorical	8	3
Education	Categorical	16	4
Country	Categorical	41	3
Marital Status	Categorical	7	3
Race	Categorical	5	2
Gender	Categorical	2	2

Group'', and ''Bank Account''. Each sensitive attribute type has 8 different attribute values under 4 distinct categories as per Table 4. First, we assign a numeric number to each sensitive attribute value under each sensitive attribute type. Second, we generate a random numeric number from 1 to 24 for each sensitive attribute values under all 3 sensitive attribute types. Then assign those numbers to the tuples/records. For example, if the random number is 5 in a tuple of the data set, then this record has the sensitive value ''Flue'' of sensitive category C, which is of sensitive attribute type ''Medical Disease''. By default, we set $\alpha = 2$, p = 4, and k = 4. Our simulation results compare our proposed (p, β) sensitive k-anonymity condition with previously proposed anonymity conditions considering bottom-up generalization in a ubiquitous network environment for service communication.

Fig. 10(a) shows the distortion ratio of our proposed (p, β) sensitive k-anonymity approach in comparison to existing anonymity approaches with respect to generalization height. Here, our (p, β) sensitive k-anonymity has a lower distortion in comparison to other existing anonymity schemes. Fig. 10(b) shows that our proposed (p, β) sensitive k-anonymity has a higher rate of tuple satisfaction in comparison to others with the increase of height.

- *Distortion Ratio:* Distortion Ratio is the ratio of required generalization height to satisfy anonymity condition and the lattice height with consideration of all tuples.

In Fig. 11(a), Pval of an entity is equal to its every trust value for any value of hop-count that greater than zero, where $W_I = 1.0$. It implies that at $W_I = 1.0$, Pval is fully dependent on one's trust, and hop-count has no importance. In Fig. 11(b), Pval of entity A with uncertain trust 0.5 and hop = h + 1 is equal to Pval of entity B with the highest trust 1.0 and hop = h. Since every trusted entity in TC_{TP} has trust value greater than uncertain trust 0.5, Pval of an entity at hop = h + 1 is always higher than an entity with hop = h, for any value of trust in our approach. It implies that, for $W_I = 0.5$, Pval is always higher for an entity with higher hop-count.

In Table 7, we have compared different parameters relevant to our work with previous privacy approaches in packet transmission during communication. Here Comp Cost, Pri type, Id Pri, Loc Pri, Bhv Pri, Untrc denotes the computation cost, privacy preservation type, identity privacy, and untraceability, respectively. All previous approaches provided user privacy either through encryption or through anonymity. In [44]–[48], privacy is provided through hash function. The computational cost of a mobile node in communications are 4Th +1Texp for [47], 8Ths for [48], 5Th for [46], 5Th for [45] and 7Th for [44]. Where 'Th' is hash computation time, and 'Texp' is Modular exponentiation computation time. These are P class problems. On the other hand, in contrast to these approaches, the k-anonymity approach is an NP hard problem [29]. All improved k-anonymity approaches such as P, P+, (P, a), (p+, a) are also NP hard problem [5], [29], [33], [34]. Along with our approach (see sec VIII). In [5], [29], [33], [34], privacy preservation mechanism is done in all time no matter whether any service communication is needed or not. But in [4], [15], [30], [34], [33] and our approach, whenever one entity wants to initiate a service communication, privacy preservation is required in an on-demand way. In all previous cases, privacy schemes are proposed to protect identity privacy. In [44]–[48] user anonymity is provided by

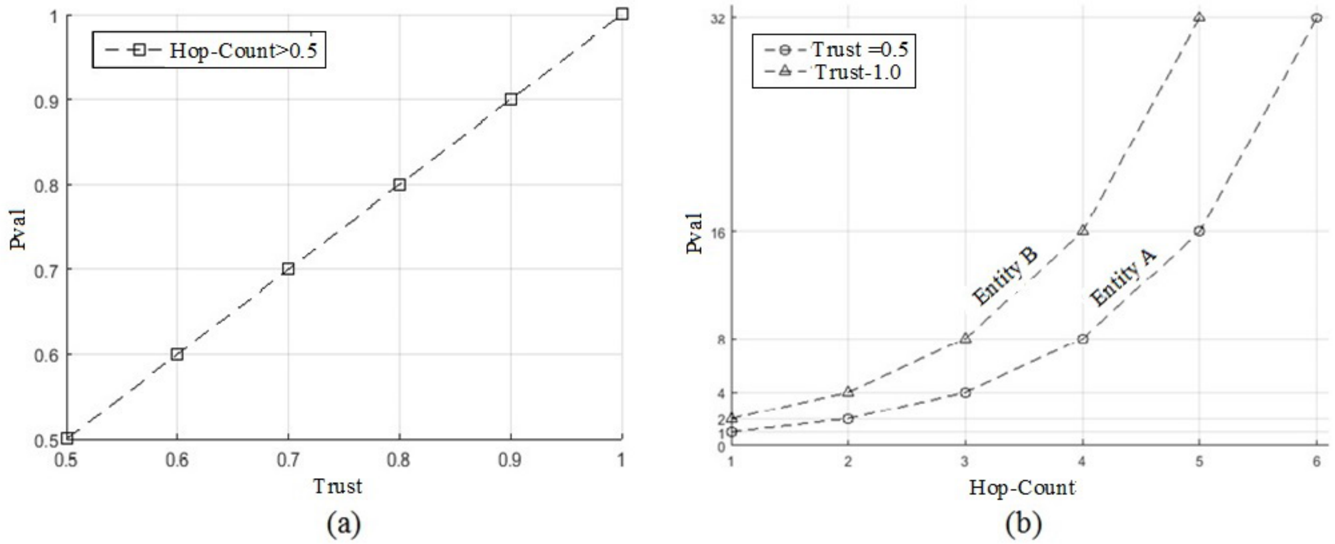

FIGURE 11. Pval vs. hop-count in our (p, β) sensitive k-anonymity approach. In Fig. 11(a), Pval of an entity is equal to its every trust value for any value of hop-count that greater than zero, where WI=1.0. In Fig. 11(b), Pval of entity A with uncertain trust 0.5 and hop=h+1 is equal to Pval of entity with the highest trust 1.0 and hop=h.

TABLE 7. Comparison table.

Algo	Comp Cost	Pri Type	Id Pri	Loc Pri	Bhv Pri	Untrc
[48]	P class	Pro-active	Yes	No	No	No
[47]	P class	Pro-active	Yes	No	No	No
[46]	P class	Pro-active	Yes	No	No	No
[45]	P class	Pro-active	Yes	No	No	No
[44]	P class	Pro-active	Yes	No	No	No
[15]	NP hard	On-demand	Yes	No	No	No
[4]	NP hard	On-demand	Yes	No	No	No
[30]	NP hard	On-demand	Yes	No	No	No
[34]	NP hard	On-demand	Yes	No	No	No
[33]	NP hard	On-demand	Yes	No	No	No
our work	NP hard	On-demand	Yes	Yes	Yes	Yes

encryption and(or) pseudo identity. In these cases, a malicious entity can reach the location of the target entity by backtracking or following the packet flow direction. On reaching the target location, a malicious entity can track the behavioural history of different pseudo identities in different sessions to create alias profiling, which leads to actual profiling of target one after some communication with similar interests. In [4], [15], [30], [34], [33], identity privacy is provided through improved k-anonymity, which is not supporting the location privacy and behavioural privacy as previously discussed approaches. Where location privacy compromised, a malicious entity may implement a jamming attack by continuously sending fake packets to that location. Sometimes, one can implement the blackhole attack by dropping all data packets coming from that location's entity. Blackhole attacker makes fool the target location's entity that it has a shorter route to destination and attracts all data packets to drop them and to disrupt the communication. If location is identified, then the identification of an entity may be possible by a continuous study of behavioural history, which may lead to the actual profiling of the identity. This may lead to impersonate attack, identity spoofing attack, and many other identity-based attacks. By tracking the behavioural history, a malicious entity may get the interests of the target one and may implement the service spoofing attack. It may communicate the target one with attractive service advertisements as per the interests. In our proposed approach, we preserved the threefold privacy in terms of not only identity but also location and behaviour.

X. CONCLUSION AND FUTURE WORK

The Privacy preservation of individual entities in the ubiquitous network for different application demands the in-depth research effort over the years and years to come. Privacy is analyzed as three-fold privacy namely, identity, location, and behaviour. A trust circle is established around the entity based on the cumulative effect of direct interactions and recommendations. The trust circle of one service entity ensures the preservation of a considered set of privacies based on direct-indirect trust and participation in 2-degree anonymity. In our 1^{st} degree anonymity, trusted identities that are "(p, β) sensitive k-anonymity" satisfied, are reported as packet sources. AlO routes that service packet acting like AcO. Consequently, a malicious entity can not divulge AcO's identity by interpreting the source of the service packet. Malicious entities fail to track AcO's behaviour due to path diversion and the anonymous behaviour of AlO. The malicious entity deviates from the actual service communicating path by our proposed 2^{nd} degree anonymity. Due to the n-deviation, a malicious entity can not follow backtracking or packet flow direction. The demand for application is fulfilled in providing personalization between anonymity and trust for AlO selection process. Our simulation with the "Facebook" data set exhibits the better performance of our proposed TC_{TP} establishment around the entities over the existing approaches. The "ADULT" data set is used to verify the proposed (p, β) sensitive k-anonymity approach, which shows better results in preserving the privacies over previous approaches. The formal model of security and privacy needs an extensive study of the existing approaches, and we are in the process of the same for a suitable formal model for privacy in the future.

REFERENCES

[1] A. L. Traud, P. J. Mucha, and M. A. Porter, "Social structure of facebook networks," 2011, *arXiv:1102.2166*. [Online]. Available: https://arxiv.org/abs/1102.2166

[2] A. Campan, T. M. Truta, and N. Cooper, "P-sensitive K-anonymity with generalization constraints," *Trans. Data Privacy*, vol. 3, pp. 65–89, Aug. 2010.

[3] S. A. Gajparia, J. C. Mitchell, and C. Y. Yeun, "The location information preference authority: Supporting user privacy in location based services," in *Proc. 9th Nordic Workshop Secure IT-Syst.*, 2004, pp. 1–6.

[4] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "L-diversity: Privacy beyond k-anonymity," *ACM Trans. Knowl. Discovery Data*, vol. 1, no. 1, pp. 106–115, Mar. 2007, doi: 10.1145/1217299.1217302.

[5] B. Zhou and J. Pei, "The k-anonymity and l-diversity approaches for privacy preservation in social networks against neighborhood attacks," *Knowl. Inf. Syst.*, vol. 28, no. 1, pp. 47–77, Jul. 2011.

[6] B. Gedik and L. Liu, "Protecting location privacy with personalized k-anonymity: architecture and algorithms," *IEEE Trans. Mobile Comput.*, vol. 7, no. 1, pp. 1–18, Jan. 2008, doi: 10.1109/tmc.2007.1062.

[7] C. Yin, J. Xi, and R. Sun, "Location privacy protection based on ImprovedK-value method in augmented reality on mobile devices," *Mobile Inf. Syst.*, vol. 2017, 2017, Art. no. 7251395.

[8] C. Zhang, X. Zhu, Y. Song, and Y. Fang, "A formal study of trust-based routing in wireless ad hoc networks," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–9.

[9] D. Xiu and Z. Liu, "A formal definition for trust in distributed systems," in *Information Security* (Lecture Notes in Computer Science), vol. 3650, J. Zhou, J. Lopez, R. H. Deng, and F. Bao, Eds. Berlin, Germany: Springer, 2005, pp. 482–489, doi: 10.1007/11556992_35.

[10] H. Chen and W. Lou, "On protecting end-to-end location privacy against local eavesdropper in wireless sensor networks," *Pervas. Mobile Comput.*, vol. 16, pp. 36–50, Jan. 2015.

[11] J. Long, M. Dong, K. Ota, and A. Liu, "Achieving source location privacy and network lifetime maximization through tree-based diversionary routing in wireless sensor networks," *IEEE Access*, vol. 2, pp. 633–651, 2014, doi: 10.1109/access.2014.2332817.

[12] J. Jiao, P. Liu, and X. Li, "A personalized privacy preserving method for publishing social network data," in *Theory and Applications of Models of Computations*, T. V. Gopal, M. Agrawal, A. Li, and S. B. Cooper, Eds. Cham, Switzerland: Springer, 2014, pp. 141–157.

[13] L. Guo, X. Zhu, C. Zhang, and Y. Fang, "A multi-hop privacy-preserving reputation scheme in online social networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2011, pp. 1–5.

[14] L. Sweeney, "Achieving k-anonymity privacy protection using generalization and suppression," *Int. J. Unc. Fuzz. Knowl. Based Syst.*, vol. 10, no. 5, pp. 571–588, Oct. 2002.

[15] L. Sweeney, "K-anonymity: A model for protecting privacy," *Int. J. Unc. Fuzz. Knowl. Based Syst.*, vol. 10, no. 5, pp. 557–570, Oct. 2002.

[16] L. Guo, C. Zhang, and Y. Fang, "A trust-based privacy-preserving friend recommendation scheme for online social networks," *IEEE Trans. Dependable Secure Comput.*, vol. 12, no. 4, pp. 413–427, Jul. 2015, doi: 10.1109/TDSC.2014.2355824.

[17] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proc. 1st Int. Conf. Mobile Syst., Appl. Services (MobiSys)*, 2003, pp. 31–42, doi: 10.1145/1066116.1189037.

[18] M. Rahimi, M. Bateni, and H. Mohammadinejad, "Extended K-anonymity model for privacy preserving on micro data," *Int. J. Commun. Netw. Inf. Secur.*, vol. 7, no. 12, pp. 42–51, Nov. 2015, doi: 10.5815/ijcnis.2015.12.05.

[19] M. S. Islam, M. A. Hamid, C. S. Hong, and B.-H. Chang, "Preserving identity privacy in wireless mesh networks," in *Proc. Int. Conf. Inf. Netw.*, Jan. 2008, pp. 1–5, doi: 10.1109/icoin.2008.4472767.

[20] N. S. Sarma and A. P. Shobak, "Friend recommendation in KNN classification," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 4, no. 6, pp. 2320–9798, 2016, doi: 10.15680/IJIRCCE.2016.0406294.

[21] P. Samarati, "Protecting respondents identities in microdata release," *IEEE Trans. Knowl. Data Eng.*, vol. 13, no. 6, pp. 1010–1027, Nov./Dec. 2001.

[22] A. Pfitzmann and M. Hansen, *Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management a Consolidated Proposal for Terminology*, document version v0.28, May 2006.

[23] P. Mayilvelkumar and M. Karthikeyan, "L-diversity on K-anonymity with external database for improving privacy preserving data publishing," *Int. J. Comput. Appl.*, vol. 54, no. 14, pp. 7–13, Sep. 2012.

[24] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, "Propagation of trust and distrust," in *Proc. 13th Conf. World Wide Web*, 2004, pp. 403–412.

[25] R. C.-W. Wong, J. Li, A. W.-C. Fu, K. Wang, "(α, k)-anonymity: An enhanced k-anonymity model for privacy preserving data publishing," in *Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2006, pp. 754–759.

[26] R. Dhekane and B. Vibber, "Talash: Friend finding in federated social networks," in *Proc. LDOW*, Mar. 2011, pp. 1–8.

[27] S. Zhang, X. Li, Z. Tan, T. Peng, and G. Wang, "A caching and spatialK-anonymity driven privacy enhancement scheme in continuous location-based services," *Future Gener. Comput. Syst.*, vol. 94, pp. 40–50, May 2019, doi: 10.1016/j.future.2018.10.053.

[28] T. Hashem and L. Kulik, "Safeguarding location privacy in wireless ad-hoc networks," in *Ubicomp 2007 : Ubiquitous Computing. UbiComp* (Lecture Notes in Computer Science), vol. 4717, J. Krumm, G. D. Abowd, A. Seneviratne, and T. Strang, Eds. Berlin, Germany: Spinger, 2007, pp. 372–390, doi: 10.1007/978-3-540-74853-3_22.

[29] T. M. Truta, A. Campan, and P. Meyer, "Generating microdata with *P*-sensitive *K*-anonymity property," in *Secure Data Management* (Lecture Notes in Computer Science), vol. 4721, W. Jonker and M. Petkovic, Eds. Berlin, Germany: Spinger, 2007, pp. 124–141, doi: 10.1007/978-3-540-75248-6_9.

[30] T. Truta and B. Vinay, "Privacy protection: P-sensitive k-anonymity property," in *Proc. 22nd Int. Conf. Data Eng. Workshops (ICDEW)*, 2006, p. 94, doi: 10.1109/icdew.2006.116.

[31] *US Census Bureau woned Adult database*. (1996). *Donor: Ronny Kohavi and Barry Becker, Data Mining and Visualization Silicon Graphics*. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/adult

[32] U. Hengartner, "Location privacy based on trusted computing and secure logging," in *Proc. 4th Int. Conf. Secur. Privacy Commun. Netw.*, 2008, pp. 22–25.

[33] X. Sun, L. Sun, and H. Wang, "Extended k-anonymity models against sensitive attribute disclosure," *Comput. Commun.*, vol. 34, no. 4, pp. 526–535, Apr. 2011.

[34] X. Sun, H. Wang, J. Li, and T. M. Truta, "Enhanced P-sensitive K-anonymity models for privacy preserving data publishing," *Trans. Data Privacy*, vol. 1, no. 2, pp. 53–66, 2008.

[35] Z. Cheng, L. Chen, R. Comley, and Q. Tang, *Identity-Based Key Agreement With Unilateral Identity Privacy Using Pairings*, vol. 3903. Berlin, Germany: Springer-Verlag, 2006, pp. 202–213.

[36] E. W. ZhWeisstein. (2006), *Bernoulli Distribution*. MathWorld—A Wolfram Web Resource. [Online]. Available: http://mathworld.wolfram.com/BernoulliDistribution.html

[37] D. Gambetta, "Can we trust trust?" in *Trust: Making and Breaking Cooperative Relations*. Oxford, U.K.: Basil Blackwell, 1990.

[38] T. Rybicki, "Semantic service discovery in pervasive computing environment," in *Proc. 5th Int. Conf. Pervas. Services*, 2008, pp. 81–90.

[39] T. Murakami and A. Fujinuma, "Ubiquitous networking: Towards a new paradigm," Nomura Res. Inst., Papers no. 2, 2000.

[40] M. A. Azad, S. Bag, S. Tabassum, and F. Hao, "Privacy preserving collaboration across multiple service providers to combat telecoms spam," *IEEE Trans. Emerg. Topics Comput.*, to be published, doi: 10.1109/TETC.2017.2771251.

[41] M. A. Azad and R. Morla, "Rapid detection of spammers through collaborative information sharing across multiple service providers," *Future Gener. Comput. Syst.*, vol. 95, pp. 841–854, Jun. 2019, doi: 10.1016/j.future.2017.12.026.

[42] M. A. Azad, S. Bag, and F. Hao, "PrivBox: Verifiable decentralized reputation system for online marketplaces," *Future Gener. Comput. Syst.*, vol. 89, pp. 44–57, Dec. 2018, doi: 10.1016/j.future.2018.05.069.

[43] Y. Zhang, R. Deng, E. Bertino, and D. Zheng, "Robust and universal seamless handover authentication in 5G HetNets," *IEEE Trans. Dependable Secure Comput.*, to be published, doi: 10.1109/TDSC.2019.2927664.

[44] H. Lee, D. Lee, J. Moon, J. Jung, D. Kang, H. Kim, and D. Won, "An improved anonymous authentication scheme for roaming in ubiquitous networks," *PLoS ONE*, vol. 13, no. 3, Mar. 2018, Art. no. e0193366, doi: 10.1371/journal.pone.0193366.

[45] S. A. Chaudhry, A. Albeshri, N. Xiong, C. Lee, and T. Shon, "A privacy preserving authentication scheme for roaming in ubiquitous networks," *Cluster Comput.*, vol. 20, no. 2, pp. 1223–1236, Jun. 2017, doi: 10.1007/s10586-017-0783-x.

[46] M. S. Farash, S. A. Chaudhry, M. Heydari, S. M. S. Sadough, S. Kumari, and M. K. Khan, "A lightweight anonymous authentication scheme for consumer roaming in ubiquitous networks with provable security," *Int. J. Commun. Syst.*, vol. 30, no. 4, p. e3019, Mar. 2017, doi: 10.1002/dac.3019.

[47] F. Wen, W. Susilo, and G. Yang, "A secure and effective anonymous user authentication scheme for roaming service in global mobility networks," *Wireless Pers. Commun.*, vol. 73, no. 3, pp. 993–1004, 2013.

[48] M. Karuppiah, S. Kumari, A. K. Das, X. Li, F. Wu, and S. Basu, "A secure lightweight authentication scheme with user anonymity for roaming service in ubiquitous networks," *Secur. Commun. Netw.*, vol. 9, no. 17, pp. 4192–4209, Nov. 2016.

[49] T. Grandison and M. Sloman, "A survey of trust in Internet applications," *IEEE Commun. Surveys Tuts.*, vol. 3, no. 4, pp. 2–16, 4th Quart., 2000.

[50] A. Josang and S. J. Knapskog, "A metric for trusted systems," in *Proc. 21st NIST-NCSC Nat. Inf. Syst. Secur. Conf.*, Arlington, VA, USA, 1998, pp. 16–29.

[51] P. Lamsal. (2001). *Understanding Trust and Security*. [Online]. Available: http://www.cs.Helsinki.FI/u/lampa/papers/UnderstandingTrustAndSecurity.pdf

[52] D. H. McKnight and N. L. Chervany, "The meanings of trust," in *Proc. Trust CyberSocieties*, vol. 2246, 2001, pp. 27–54.

[53] E. C. Tomlinson and R. J. Lewicki. (2002). *Trust and Trust Building*. [Online]. Available: http://www.beyondintractability.org/m/trustbuilding.jsp

SWARNALI HAZRA received the M.Tech. degree from the University of Calcutta. She is a Research Scholar with the Department of Computer Science and Engineering, University of Calcutta. Her research interests include trust, network security, privacy, network overhead, and so on.

S. K. SETUA is an Associate Professor with the Department of Computer Science and Engineering, University of Calcutta. He has published more than 50 research articles in peer-reviewed journals and conferences. His research interests include security, privacy, parallel, and distributed computing, image processing, software-defined systems, and so on.