Addressing False Identity Attacks in Action-based P2P Social Networks with an Open Census

S. Qin*, Marius C. Silaghi*, T. Matsui†, M. Yokoo‡ and K. Hirayama§
*Florida Tech
†Nagoya Inst. of Technology
‡Kyushu Univ.
§Kobe Marine Univ.

***Abstract*—P2P social networks defined by user actions (e.g., P2P discussion forums) are expected to be ideal environments for Sybil and false identity attacks (just as in the case of the similar web based systems: YouTube, etc.). In particular, these attacks are a significant impediment for meaningful electronic petition drives since they render impossible the verification of the eligibility of participants.**

While many electronic social networks strive for guaranteeing the privacy (e.g., by anonymization) of their users, existing systems for petition drives, like DirectDemocracyP2P.net, encourage users to disclose their real identities and are meaningless when users do not follow this request.

We describe a framework and investigate techniques for running decentralized peer-to-peer census processes that enable observers to independently verify the identity of participants in a social network.

I. Introduction

We address the problem of detecting false identities in P2P systems for petition drives by gathering census data using a decentralized, citizen-driven mechanism. While digital certification from governments could offer an alternative to solution, most governments are shy to offer such certificates, and moreover it would also require trust in these governments (trust that certain governments could use to silence opposing voices by not delivering them the needed certificates). The challenge addressed here consists in formalizing the census problem and developing algorithms applicable in a peer-to-peer (P2P) approach. The result of the performed census is expressed in a trust value for the reliability and eligibility of each identity, value that integrated in processes such as petition drives, debates and polls targeting well defined populations.

After introducing related work, in section Concepts we introduce the main definitions. Section Techniques introduces the experimented algorithms for evaluating user data. We conclude after discussing experimental results.

II. Background

One of the main challenges of large distributed collaborations is that one user can login under as many identities as she has time and desire to register. The creation and usage of such duplicated identities is referred in literature as the Sybil attack. The term Sybil attack was first introduced by [1] in a generic distributed computing environment. In the presence of a trusted authority, the resistance to Sybil attacks is either offered by explicitly certified participation as in Microsoft's Farsite [2]

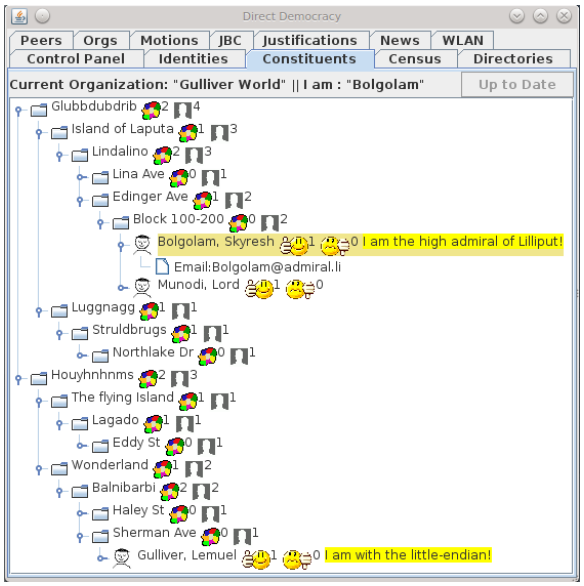

Fig. 1. Tree Structure of Neighborhoods (Locations from Gulliver's Travel).

or by an implicit verification. This implicit verification can be regarded as too dependent on unsafe assumptions about underlying systems, as in the Cooperative File System [3].

The concept of regional/neighbor based trust and verification is used in the Thawte Web of Trust [2]. There, local trusted people called notaries can verify one's credentials and certify them using a Thawte certificate. Regional/neighbor based trust and verification is also used with PGP, where people meet for key signing parties, giving each other an independent proof of identity after manually inspecting government issued documents.

Census processes with validation can be successful only if people are sufficiently connected to provide enough data to the decision making process. Studies of connectivity between people have been conducted in relation to existing social networks. A kind of constituency was discussed in [4].

A reputation system maintains scores inferred from other's opinions for participants [5], [6]. Notions of valued trust are proposed in [7] and extended in [8]. The introduced values can be used to decide if an entity is sufficiently trustworthy. The values are inferred from a graph with nodes as entities and edges as the trust relations. They formalize trust relations of different types, among which are identification (ID) and trust-

worthiness (PR), and discuss the potential offered by networks of such relations to model known distributed authentication protocols. Solutions based on Dynamic Belief Networks are proposed in [9].

III. Census Process Concepts

The concepts of organization, neighborhood, constituent and witnessing are introduced in [9]. Now let us introduce concepts involved in the decentralized census processes.

A. Citizen Interactions

A citizen-driven census requires participation of individual citizens for actions such as *residence declaration* and *witnessing*. As residence declarations, each individual voluntarily provides census data not only about herself but also about her neighbors. The neighborhood where a citizen resides is part of its identity details.

a) Verification: A voting process, called witnessing, is used to help verify the census data. The verification can be done both by neighbors, and by volunteers who gather data about the inhabitants of the given area.

b) Witness Graph: A graph defined by the witness relations between constituents can be generated in the following way:

- A node is generated for each constituent.
- A directed edge from node A to node B is generated for each semantic statement that A witnesses for B.
- Each edge has a color (from a set Ω), given by the type of statement that generated it (ontological commitment).
- An edge has weight 1 if generated for a favorable stance and weight 0 if generated for an unfavorable stance (epistemological commitment).

Inactive nodes are sinks for this graph. This graph can be used to reason about the eligibility of the declared identities and implicitly about the census.

c) Distributed Census Problem: The Distributed Census Problem (DCP) for an observer Γ can be formalized as a tuple $\langle \mathcal{N}_\mathcal{S}, \mathcal{I}, \mathcal{R}, \mathcal{W}, \mathcal{M}_\mathcal{S}, \Gamma \rangle$, where:

- $\mathcal{N}_\mathcal{S}$ is the set of neighborhoods $\mathcal{N}_\mathcal{S} = \{1, ..., d\}$,
- $\mathcal{I}$ is the set of person identities,
- $\mathcal{R}$ is the set of residence declarations (constituent items)
- $\mathcal{W}$ is the set of witness stances
- $\mathcal{M}_\mathcal{S}$ is a model of the relation between the ground truth I^* and $\mathcal{N}_\mathcal{S}$, $\mathcal{I}$, $\mathcal{R}$ and $\mathcal{W}$, as believed by the observer Γ (e.g., a certain belief network)

$\mathcal{I}^*$ (the ground truth), each having an identity from the set $\mathcal{I}$. The problem is to approximate the $\mathcal{I}^*$ that best explains $\mathcal{N}_\mathcal{S}$, $\mathcal{I}$, $\mathcal{R}$ and $\mathcal{W}$ based on the model $\mathcal{M}_\mathcal{S}$.

IV. Techniques

Here we present the techniques used to address the challenge of inferring a count of the constituency given a witness graph.

d) Eligibility: Although anyone can participate in the census process of an organization, not everyone is eligible to be counted in the census. In an organization, which is the context of this study, the definition of eligibility is a function of the constituent. When the eligibility for a constituent is based on a subjective view, the census result is relevant only to the user (or users) sharing this view. Hence, we define the eligibility as a probabilistic function of several parameters:

- Someone's interpretation of the witness graph, $\mathcal{M}_\mathcal{S}$
- Someone's own definition of the eligibility, $\Gamma(\mathcal{O})$

Definition 1: The reference user is the user Γ who currently computes the census.

Definition 2 (Censable and Ψ): A constituent item C is `censable` for an organization if it is eligible and new (never counted elsewhere). The Γ's confidence value in whether C is `censable` is denoted $\Psi(C)$.

Definition 3 (Witness Reliability and Φ): A constituent item C is a *reliable witness* if Γ trusts all the witness stances that C issues as she trusts her own. Γ may not fully trust the stances of another constituent C, but only with a confidence value $\Phi(C)$.

Based on the DCP parameters, one can infer a value Ψ for the confidence that observer Γ can have on whether a given constituent item C identifies a `censable` user, and a value Φ for its confidence on whether C is *witness reliable*.

Remark 1 (Decision Criteria 1): One approach to compute a census is to declare that an identity is eligible (to be counted in the census) from the point of view of the reference user Γ if the value of Ψ surpasses a threshold t where t is defined by the Γ.

Remark 2 (Decision Criteria 2): Another approach is to sum the values Ψ for all constituents (once normalized in the interval [0,1]).

In the next paragraphs, we proposed algorithms to compute the Ψ value for each node in the witness graph for various types of models $\mathcal{M}_\mathcal{S}$.

For a given type of semantic statement, we will use the following concepts:

- The $supporting\ parents$ of the node C that witness C favorably for the quality q, $q \in \Omega$, are denoted as $SP^q(C)$.
- The $opposing\ parents$ of the node C that witness C unfavorably for the quality q, $q \in \Omega$, are denoted as $OP^q(C)$.
- The $supported\ children$ of the node C are the children that are witnessed favorably by C for the quality q, $q \in \Omega$, denoted as $SC^q(C)$.
- The $opposed\ children$ of the node C are the children that are witnessed unfavorably by C for the quality q, $q \in \Omega$, denoted as $OC^q(C)$.

- The *amortization factor* $f_q, f_q \in [0,1]$, models the decrease of the confidence during transfer by witnessing for quality q. These factors compose as one gets further from Γ in the transitive chain of trust (along the *"reliable witness"* edges in the witness graph).

We introduce the notation Φ to denote the quality *reliable witness* when used as superscript or subscript with one of the notations above (e.g., $SC^{\Phi}(C)$). Similarly we use Ψ to denote the `censable` quality when used as superscript or subscript in these notations (e.g., $SC^{\Psi}(C)$).

Example 1: Given a supporting parent node sp of the node C, the confidence $i_{sp,C}$ propagated (in certain introduced models) from sp to C is $\Phi(sp) \times f_{\Phi}$.

Assume that a *supporting parent* sp of node C has value $\Phi(sp)$=0.8 and Γ's f_{Φ} is 0.9, the confidence $i_{sp,C}$ transferred from sp to C is 0.8×0.9=0.72.

Remark 3: However, amortization may not apply to opposing parents even as it applies for supporting parents. In certain introduced models, given an opposing parent node *op* of the node C, the confidence $i_{op,C}$ propagated from *op* to C is 0.

The node representing the reference user Γ and its directly connected children are treated separately and referred to as *special nodes*.

Several of the proposed approaches share the following assumptions for the value of the *special nodes*:

Assumption `[Self_Trust]`

- $\Psi(a) = 1, \forall a \in SC^{\Psi}(\Gamma)$
- $\Psi(a) = 0, \forall a \in OC^{\Psi}(\Gamma)$
- $\Phi(a) = 1, \forall a \in SC^{\Phi}(\Gamma)$
- $\Phi(a) = 0, \forall a \in OC^{\Phi}(\Gamma)$

The models of approximate reasoning introduced next are: Max Amortized Support (MAXAS), Adjusted Max Amortized Support (AMAS), Average Support (AS), Penalized Average Support (PAS), and Adjusted Support Ratio (ASR).

e) Maximum Amortized Support (MAXAS): The first model we introduce for computing the Ψ value of each node in the witness graph is given in Algorithm 1. This model employs the amortization factor (Line 11) as per Example 1. The values of f_{Φ} and f_{Ψ} are user provided inputs to the algorithm. For each node, the initial Ψ and Φ are 0 since Γ a priori knows nothing about it (Line 1).

The *reference user*, Γ, who computes the census has full confidence in her witness stances (Lines 2 and 3), as per the assumption `Self_Trust`. Further, the algorithm traverses remaining nodes in the graph in breadth first order. If a node C has N *supporting parents*, $SP^{\Phi}(C)$, MAXAS computes $\Psi(C)$ as the maximum out of all the confidence values transferred from them (Line 11).

An example is given in Figure 2, and we use it next to illustrate the output of Algorithm 1. In this figure, for simplicity, a single edge is used to represent all semantic statements in a witness stance, and all of them are supposed to have the same weight (epistemological commitment, 1 or

Input: A witness graph g with the starting node Γ and amortization factor f_{Φ} for the *reliable witness* quality and f_{Ψ} for the `censable` quality

```
1  for each node n in g do Ψ(n) ← 0; Φ(n) ← 0;
2  Ψ(a) ← 1, ∀a ∈ SC^Ψ(Γ); Ψ(a) ← 0, ∀a ∈ OC^Ψ(Γ);
3  Φ(a) ← 1, ∀a ∈ SC^Φ(Γ); Φ(a) ← 0, ∀a ∈ OC^Φ(Γ);
4  Add SC^Φ(Γ) to queue Q;
5  while Q is not empty do
6      node n ← extract_first(Q);
7      foreach c ∈ SC^Ψ(n) do
8          Ψ(c) ← max(Ψ(c), f_Ψ × Φ(n))
9      end
10     foreach c ∈ SC^Φ(n) do
11         Φ(c) ← max(Φ(c), f_Φ × Φ(n));
12         if c has never been added to Q then
13             add c to the end of the Q
14         end
15     end
16 end
```

Algorithm 1: Derivation of $\Psi(C)$ and $\Phi(C)$ for each constituent C

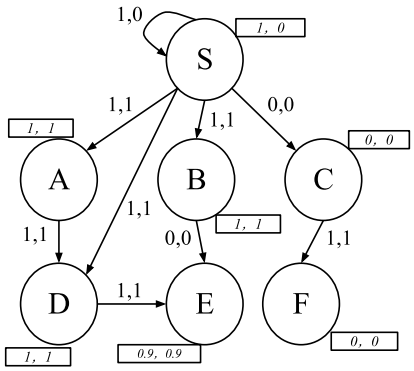

Fig. 2. Propagation of (Φ, Ψ) value pairs

0). The amortization factors are also equal: $f_{\Phi} = f_{\Psi} = 0.9$. Initially, each node's Ψ value is set to 0 (Line 1). Node S is the *reference user* Γ and the shown graph captures her data about other nodes (A, B, C, etc). The nodes are labeled with the confidence in their *witness reliability* that S infers from this graph. One can see that she trusts herself ($\Phi(S) = 1$, Line 2), and does not considers itself `censable`, ($\Psi(S) = 0$). S also trusts those for whom she issues stances as favorable *reliable witness* (A, B and D, $\Phi(A) = \Psi(A)$=1, $\Phi(B) = \Psi(B)$=1, $\Phi(D) = \Psi(D)$=1, Line 2). A, B and D are added to queue Q in this order. For node C, since $\Phi(C)$=0 we stop the distribution of trust to C's children, hence do not infer anything about F ($\Psi(F)$ and $\Phi(F)$ remain 0). A is then dequeued from the head of the queue (Line 6). Since A has a favorable witness on D, now $\Phi(D)$ is evaluated to be $max(1, 0.9)$=1 (Line 11). We do not add D to the queue, since it is already in it. Then B is dequeued, but it has no favorable witness on its only child E. $\Psi(E)$ is not changed. Then D is dequeued. Since D has a favorable witness on E, $\Psi(E) = \Phi(E)$=$max(0.9, 0)$=0.9 (Line 11) and E is added to Q. C and F are never added to the queue.

In the context of Algorithm 1, and when the census is

estimated with the mechanism in Remark 1, in order to be counted, a node has to be within a certain support distance from the root s. The distance is determined by the used amortization factor and threshold. The smaller the factor, the smaller the required distance. In the example given by Figure 2, assume we choose the census threshold t as 0.95, S, A, B and D are counted. Note that E is not counted because its support path exceeds the distance of 1=1 + $\lfloor \log_f t \rfloor$ edges from S.

f) Adjusted Max Amortized Support (AMAS): In this second model, to enable the increase of Ψ for a constituent when it gets extra support, we let $\Psi(C)$ take values between $M(C)$ and $N(C)$ ($\Psi(C) \in [M(C), N(C)]$) where:

$$M(C) = \max_{n \in SP^{\Psi}(C)} \Phi(n) \times f_{\Psi}$$

$$N(C) = \max_{n \in SP^{\Psi}(C)} \Phi(n)$$

The algorithm we use to compute the Φ value here is similar to Algorithm 1. We do not repeat the algorithm and only address the difference. In Algorithm 1, the Ψ value of a constituent item C is computed as $M(C)$. This model assumes Equation 1:

$$\Psi(C) = M(C) + \frac{(N(C) - M(C)) \times min(fw, W)}{W} \quad (1)$$

In Equation 1, fw is the total number of favorable `censable` witnesses ($|SP^{\Psi}(C)|$) for C and W is a user-defined parameter. The closer W grows towards $|SP^{\Psi}(C)|$, the closer $\Psi(C)$ approaches to $M(C)$. If C has more than W favorable witnesses from its parents, $\Psi(C)$ becomes $M(C)$.

g) Average Support (AS): In the first two approaches, note that $\Psi(C)$ is only inferred from the *supporting parents*. For the *opposing parents*, the propagation is stopped, i.e., the inputs from the *opposing parents* to the children nodes are 0s. In this approach and the next one, we employ the inputs from the *opposing parents* into computing the Ψ value of a node in the witness graph. AS can be seen as an extension with amortization factors of an interpretation of the method to combine recommendation trust proposed in [8], where recommendation values are replaced by the confidence in the witness reliability of constituents witnessing an entity. .

The AS model computes the Ψ value using Equation 2 where $|SP^{\Psi}(C)|$ is the number of favorable witnesses for C and $|OP^{\Psi}(C)|$ is the number of unfavorable witnesses for C.

$$\Psi(C) = \frac{f_{\Psi} \times \max\left(0, \sum_{n \in SP^{\Psi}(C)} \Phi(n) - \sum_{n \in OP^{\Psi}(C)} \Phi(n)\right)}{|SP^{\Psi}(C)| + |OP^{\Psi}(C)|} \quad (2)$$

The trust associated with a constituent item is also computed using a similar expression:

$$\Phi(C) = \frac{f_{\Phi} \times \max\left(0, \sum_{n \in SP^{\Phi}(C)} \Phi(n) - \sum_{n \in OP^{\Phi}(C)} \Phi(n)\right)}{|SP^{\Phi}(C)| + |OP^{\Phi}(C)|} \quad (3)$$

$Constituent$	$\Psi(C)$	Counted Constituent
C_1	2.75	$Passed$
C_2	2.0	
C_3	2.0	
C_4	4.0	$Passed$
...	...	...

TABLE I. SAMPLE OUTCOME WITH MODEL ASR FOR S_w=6, O_w=4, t=2

h) Penalized Average Support (PAS): In the fourth discussed model, we build on the AS model but reduce the penalty introduced by opposing parents to only the part in the denominator of the fraction.

$$\Psi(c) = \frac{f_{\Psi} \sum_{n \in SP^{\Psi}(C)} \Phi(n)}{|SP^{\Psi}(C)| + 1 + f_{\Psi} \sum_{n \in OP^{\Psi}(C)} \Phi(n)} \quad (4)$$

The Φ value is also computed with a similar expression, just using the SP^{Φ}, OP^{Φ} and f_{Φ}, instead of the corresponding values for Ψ in Equation 4.

i) Adjusted Support Ratio (ASR): This last non-probabilistic approximate model is much simpler than previous ones. The heuristic used here is that the identity of a constituent is more likely to be considered `censable` by others if she has relatively more SP^{Ψ} than OP^{Ψ}. An equation reflecting this value is:

$$\Psi(C) = \frac{|SP^{\Psi}(C)|}{|OP^{\Psi}(C)|} \quad (5)$$

As computed in Equation 5, $\Psi(C)$ is respecting the heuristic in the sense that a larger SP^{Ψ} indicates a larger $\Psi(C)$ and well as a smaller OP^{Ψ}. However, division by zero has to be avoided, and the obtained range of $\Psi(C)$ may need to be adjusted for a given data. Hence S_w and O_w are added as user specified parameter to allow users to adjust the range of $\Psi(C)$. A sample output of the census process using Decision Criteria 1 is shown in Table I.

$$\Psi(C) = \frac{|SP^{\Psi}(C)| + S_w}{|OP^{\Psi}(C)| + O_w} \quad (6)$$

V. EXPERIMENTS

To evaluate the power of the studied DCP models to represent users reasoning about census, as well as to study the impact of interactions between constituents , we perform a sets of preliminary experiments based on a set of volunteers.

In the experiments based on volunteers we asked 10 people living within an area of a few square kilometers to register themselves as active constituents and to also register others 10 friends as inactive constituents of a regional organization. Each of these volunteers had the opportunity to witness for the other constituent items that they knew. We also introduced 2 obviously wrong constituents at an address that most participants knew to not exist. A snapshot of the interactions between constituents is shown in Figure 3 where the thick edges represent favorable witness stances, the thin edges represents unfavorable witness stances, the nodes represents constituents and size of node is proportional to the in-degree. Red nodes denote active constituents and blue nodes denote inactive ones.

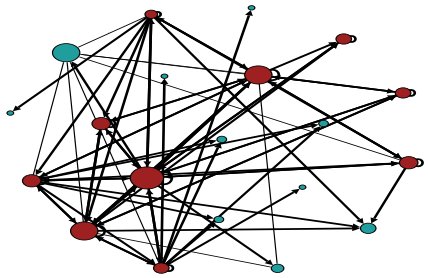

Fig. 3. Visualization of constituents and the witness relation between them

After the P2P witnessing process reached quiescence, we asked 5 of the participants (who were available) to use the widgets implementing each of the available five non-probabilistic DCP models for deciding on a census based on the available constituent items. Each of these constituents ranked the five models in terms of how well they were able to capture their own opinion on the `censable` status of each item and on their correctness. A score between 0 and 10 was assigned to each model. These scores are detailed in the Table II. While the size of the sample is small and deviation of these scores is high, currently the winner is the Penalized Average Support model (PAS). It is remarkable that the ASR, which is a very simple computation performed also acceptably well.

Constituent	C_1	C_2	C_3	C_4	C_5	Average
MAXAS	8.5	5	10	3	7	6.7
AMAS	9.5	3	10	5	7	6.9
AS	8.5	5	10	6.5	7	7.4
PAS	9	8	10	6	7	8
ASR	10	2	6	8	9	7

TABLE II. MODELS AND THEIR SCORES. THE PREFERRED PARAMETERS FOR MAX ARE $t = 0.5$, $f = 0.5$, FOR AS ARE $t = 0.65$, $f = 0.7$, FOR PAS ARE $t = 0.43$, $f = 0.82$, FOR ASR ARE $|S_w| = 6$, $|O_w| = 4$, AND $t = 2$.

VI. CONCLUSIONS

We address the problem of detecting false identities in action-based P2P social networks. The handling of these false identities is particularly important for the verification of eligibility in applications such as petition drives. Given the common assumptions of correct revelation of identities for these application, we observe that a census process of the social network base can be employed to mitigate this problem. While the real census and certification of certain countries could be contributed by governments, that is not possible for all organization, and legal issues could stop others from helping. Moreover, while population census is an important process with large implications in the distribution of public funds and security of elections from vote stuffing, it is currently an expensive process outside the reach of external verifiers and was identified as a threat to stability in certain regions.

To enable a decentralized citizen-driven population census, we employ a set of concepts such as: grassroot organization, constituent, witnessing. The grassroot organization is a set of rules (constitution) that specifies mechanisms to define eligibility of constituents. Constituents can witness (vote) on each other's qualities, such as: eligibility and witnessing reliability.

The concept of neighborhood is introduced and formalized in order to improve the scalability of peer verification. Neighborhoods group constituent addresses in a tree structure. For large organizations, the constituency is organized in a tree of neighborhoods to help with census organization. Census results can be verified separately for each neighborhood and it is reasonable to expect most users to be able to verify the existence of the immediate child neighborhoods of all the nodes on the path from their own address to the root. This enables a distributed verification of the existence of declared neighborhoods (and thereby of addresses).

Items for these concepts are identified by global identifiers guaranteed to be unique and that are disseminated among peers based on P2P protocols (current experiments being based on the DirectDemocracyP2P platform). A peer is a user acting under one name and public key via multiple agents (e.g., one agent per device that she uses).

A set of five efficient but approximate models of relations between witness stances and properties of constituents are proposed and empirically evaluated. We have also proposed and analyzed theoretically a probabilistic model based on Bayesian Networks that can be used to address the problem in a principled way [9]. Markov Chain Monte Carlo inferences are found to converge within few seconds for reasonable sized neighborhoods, and scale linearly with the input size. Preliminary experiments with volunteers are used to rank these models, while experiments with large simulated data show that robustness to attackers is possible when there exists a reasonable kernel of honest active constituents.

REFERENCES

[1] J. Douceur, "The sybil attack," *Peer-to-peer Systems*, pp. 251–260, 2002.

[2] Thawte, "Web of trust," http://www.thawte.com/secure-email/web-of-trust-wot/, 2009.

[3] F. Dabek, M. Kaashoek, D. Karger, R. Morris, and I. Stoica, "Wide-area cooperative storage with CFS," in *Proceedings of the eighteenth ACM symposium on Operating systems principles*. ACM, 2001, pp. 202–215.

[4] J. Fedoruk, A. ; Denzinger, "A general framework for multi-agent search with individual and global goals: Stakeholder search," *International Transactions on Systems Science and Applications (ITSSA)*, vol. 1, no. 4, pp. 357–362, 2006.

[5] P. Resnick, K. Kuwabara, R. Zeckhauser, and E. Friedman, "Reputation systems," *Communications of the ACM*, vol. 43, no. 12, pp. 45–48, 2000.

[6] M. U. Zeinab Noorian, "The state of the art in trust and reputation systems: A framework for comparison." *JTAER*, vol. 5, no. 2, pp. 97–117, 2010.

[7] R. Yahalom, B. Klein, and T. Beth, "Trust relationships in secure systems-a distributed authentication perspective," in *Research in Security and Privacy, 1993. Proceedings., 1993 IEEE Computer Society Symposium on*. IEEE, 1993, pp. 150–164.

[8] T. Beth, M. Borcherding, and B. Klein, "Valuation of trust in open networks," *Computer Security-ESORICS 94*, pp. 1–18, 1994.

[9] S. Qin, M. C. Silaghi, T. Matsui, M. Yokoo, and K. Hirayama, "Reputation system for decentralized population census," in *Proocedings of WIT-EC*, 2013.