

Norm's Trust Model to Evaluate Norms Benefit Awareness for Norm Adoption in an Open Agent Community

Al-Mutazbellah Khamees Itaiwi
College of Graduate Studies
Universiti Tenaga Nasional
Kajang, Selangor, Malaysia

Mohd Sharifuddin Ahmad, Alicia Y. C. Tang
College of Computer Science &
Information Technology
Universiti Tenaga Nasional
Kajang, Selangor, Malaysia

Abstract—In recent developments, norms have become important entities that are considered in agent-based systems' designs. Norms are not only able to organize and coordinate the actions and behaviour of agents but have a direct impact on the achievement of agents' goals. Consequently, an agent in a multi-agent system requires a mechanism that detects specific norms for adoption while rejecting others. The impact of such norms selection imposes risks on the agent's goal and its plan ensuing from the probability of positive or negative outcomes when the agent adopts or reject some norms. In an earlier work, this predicament is resolved by enabling an agent to evaluate a norm's benefits if it decides to adopt a particular norm. The evaluation mechanism entails a framework that analyzes a norm's adoption ratio, yield, morality and trust, the unified values of which indicates the norm's benefits. In this paper, the trust parameter of the mechanism is analyzed and a norm's trust model is proposed and utilized in the evaluation of a norm's benefits for subsequent adoption or rejection. Ultimately, the norm's benefits are determined as a consequence of a favorable or unfavorable trust value as a significant parameter in a norm's adoption or rejection.

Keywords—Norm's benefits; norm's trust; norm detection; normative multi-agent systems; intelligent software agent

I. INTRODUCTION

Trust is one of the most important aspects in human relations. In its absence, we face problems with those around us, because trust is the basis of relations in all its forms. There are many connotations of trust in a social context [1]. Thus, trust is defined as a relationship of dependence between two parties; the first party (trustor) has the confidence to rely on another party (trustee) to adopt its actions [2], [3]. Therefore, relationships between people can be inferred from trust. Conceptually, trust is also referred to relationships within and between social groups (families, friends, communities, organizations, companies, nations, etc.). It is a popular approach to frame the dynamics of group interactions in terms of trust [4].

In sociology and psychology, trust is the subject of continuous research to measure the degree of trust to another, which is the extent of belief in honesty from the other party. According to Romano [5] who views trust from the standpoint of multiple disciplines, "trust is a subjective assessment of

another's influence in terms of the extent of one's perceptions about the first-rate and significance of another's influence on one's consequences in a given situation, such that one's expectation of, openness to, and inclination towards such influence grant a sense manage over the achievable outcomes of the situation".

Trust can be seen as betting on potential contracts, which may bring benefits. Once the bet has been determined (i.e., confers trust), the trustor suspends his/her disbelief and does not consider the possibility of taking any negative action at all. Because of this, trust acts as a redactor of social complexity [6]. This phenomenon [7] can be compared with studies on social actors and their decision-making process, in the expectation that the understanding of this process (and modelling) permits the emergence of trust. Therefore, trust is part of the idea of social influence and on this basis, trust can be seen as a personal trait that increases personal relationships.

In an earlier work [8], it is proposed that intelligent agents should adopt or reject norms based on their awareness of the norms' expected benefits or losses rather than by sanctions or imitating other agents. Consequently, a framework constituting agents' awareness of norms' benefits is proposed, which is a formulation of Norm's Adoption Ratio, Yield, Trust, and Morality. With these parameters, agents compute the benefits of detected norms and subsequently determine whether the norms increase or decrease their utilities for eventual adoption or rejection.

Norm's Trust (NT) is one parameter in the formulation that motivates an agent to adopt a norm when the agent is able to compute a norm's trust value. A norm's trust refers to the degree of an agent's belief in a norm that influences other agents to adopt the norm. If the trust value of a particular norm is high, it increases the possibility of adopting the norm.

The motivation in this work stems from the need for software agents to detect and recognize the norms that are prevailing in a society of agents. In open normative-MAS, agents adopt norms to increase their utilities.

Implementations for such adoption are manifested by mechanisms, which are based on sanction, imitation, or social learning. However, without analyzing these norms, agents

ultimately adopt the norms, 'unaware' of its benefits for its adoption. However, in real world situations, a number of agents persistently violate the norms for their benefits, which may offer advantages in its quest to achieve their goals. Hence, it is proposed, in this work, that intelligent agents should adopt norms based on their 'awareness' of the norms' expected benefits on their utilities and not merely by sanctions or imitating other agents.

In open-MAS, numerous types of norms are enacted in many multi-agent societies. Consequently, a visitor agent must be able to evaluate all norm variations in these societies. To avoid the adverse effect of failure to comply with a society's norm, an agent must be able to evaluate a norm's trust, which is one of the factors that is perceived as beneficial for the agent in achieving its goals [9].

In this paper, the work-in-progress of the research in norm's benefits awareness is presented. It discusses the final parameter in formulating a norm's benefit, which is the norm's trust. The paper is organized as follows: Section II reviews the literature in this area. Section III discusses the development process. Section IV introduces the concept of a norm's benefit. Section V explains the concept of norm's trust. Section VI discusses the evaluation of the norm's trust. Sections VII and VIII present the social simulation and Section IX concludes the paper.

II. LITERATURE REVIEW

Norms are essential for the conduct of a society to establish order and harmony. Generally, people exercise the norms when they are in a new society, and occasionally, violations of the norms may be subjected to punishment or rejection by the community [10], [11]. Conversely, rewards are conferred in some cases of norms compliance. For example, when we are in a foreign country and want to use a train, we may notice people queuing, sitting and loitering while waiting for the arrival of the train. It comes to mind whether the norm (queuing, sitting or loitering) is trusted or avoiding it will lead to the failure to embark the train and have to wait for the next train? [7]. In this case, it is possible to rely on certain sources to ascertain the trustworthiness of this norm. One of these sources is to enquire the authorized people at the station about that norm and whether it is trusted or distrusted [7], [12].

Occasionally, we need to know information about some things in our society and usually, we ask competent authorities. For example, if we want to know the difference between Einstein's General Theories and Special Theories of Relativity, we will certainly ask people with a specialty in Physics. This is also the case if we want to know a trusted norm in a society and how trustable it is to apply it in that environment. It is better to ask information from the authorized people in that environment. Van Dijke shows in his study how an authority affects the behaviour of workers and increase their trust in high-level authority [13].

Another reliable source is the reputation of a norm. For example, if we are looking for a new dishwasher, we would probably pick up a copy of the Consumer Report, or we may ask our friends or neighbors if they are happy with a particular

brand and that would help us to choose the right one. Similarly, we use the reputation of a norm if we do not have sufficient information as to whether or not the prevailing norm is trusted. In the context of the Semantic Web, Van Dijke et al. shows an overview about the difference between the reputation metrics and explains that the reputation metrics are of two types, which are global and local reputation metrics [13]. Kiefhaber et al. shows that an entity can ask their neighbors about the reputation of another entity, their opinion of the target entity that will get transferred to their neighbors and so on [14].

Many scholars differ in their definitions of the concept of trust. Some define trust as part of the social and cognitive aspects of an organization, and many of the literature refers to it as one of the most important components of society [5], [15], [16]. Trust is an interactive relationship and a complex organizational structure between two or more parties. It arises from the urgent need to interact with members of a community. This relationship requires reliance on the others to achieve a specific goal. To establish this trust, the relationship between the parties must be free from anxiety. It is to trust or rely on someone's ability or involvement.

III. DISCUSSION

The literature provides useful information for the development and computation of the norm's benefits concept that incorporates trust as a computational element. Topics in norms, norms detection, trust and reputation are reviewed, which provide general and basic ideas that are important to build the trust model.

While there are many techniques of norms detection that have been proposed by researchers, the issue of open MAS has made the problem somewhat complex when dealing with similar norms in multi-agent societies. Consequently, the concept of norm's benefits is chosen to enable agents to compute specific factors that contribute to the objective determination of norms for adoption in these societies.

IV. CONCEPT OF NORM'S BENEFITS

The parameters that constitute the norms' benefits are identified from the review and analysis of the literature. In a previous work [8], these parameters are proposed to include the Norm's Adoption Ratio, Norm's Yield, Norm's Morality, and Norm's Trust. The significance of these parameters is justified by assessing the influence of each of the parameter on the decision of agents to adopt or reject a norm:

- **Norm's Adoption Ratio (N_{AR}):** It is the ratio of agents enacting a particular norm to the population of agents in a community. If P is the agents' population, and N_a is the number of agents enacting a particular norm, then $N_{AR} = N_a:P$. A high ratio is obtained when a majority of agents enact a norm while experiencing its benefits. Such experience reinforces an agent's decision to enact the norm and gain the expected benefits or violate the norm to avoid expected losses. For example, in an elevator scenario, if a majority practices the norm of *excusing* oneself when exiting the

elevator, an agent expects that the benefit from adopting such norm increases its reputation.

- **Norm's Yield (Ny):** A norm's yield is the expected gain received from adopting a norm arising from the norm's return on an agent's utility. When an agent discovers the yield of a particular norm, it infers the benefits of adopting the norm. If the norm possesses high yield, it motivates the agent to adopt it. For example, reading news online becomes the norm of many communities because it is inexpensive and convenient.
- **Norm's Morality (Nm):** This refers to the state of a norm (good or bad) with reference to a moral code. The morality of a norm allows an agent to check whether the norm conforms to its moral code. If it conforms, the probability of adopting the norm is high and vice versa. For example, talking loudly or shouting is generally considered as a low morality norm for many communities. But if it is computed as a strong norm in a particular community, an agent has the option to accept or reject the norm basing on the norm's expected benefits.
- **Norm's Trust (Nt):** A norm's trust refers to the degree of an agent's belief in a norm that influences other agents to adopt the norm. If the trust value of a particular norm is high, it increases the possibility of adopting the norm. Andrighetto et al. [17] exemplify a bus stop scenario of a particular community, in which when people arrive at the bus stop, they do not form a queue but sit on a bench and memorize who came earlier than them. In such situation, because people highly trust the norm, they adopt the norm.

If an agent is able to determine the values of the above parameters, it can compute the norm's benefits, which offers a more elegant method to adopt or reject the norm.

Fig. 1 shows a proposed norm's benefits model. A visitor agent observes and evaluates the parameters' values (i.e., Norm's Adoption Ratio, Norm's Yield, Norm's Trust, and Norm's Morality). Having determined the parameters' values, e.g. high; medium; or low, the agent's belief is influenced by these values, which in turn influence its decision to adopt or ignore the norm.

V. CONCEPT OF NORM'S TRUST

Norm Trust, as a research topic, has several meanings. For example, McKnight and Chervany [2] refer trust to one party who is willing to rely on the actions of another party. For the purpose of this research:

Definition 1: A Norm's Trust is the degree to which an agent can be expected to rely on the social norms that are believed, applied and followed without adversely affecting its objectives while reaping the norm's benefits.

A. The Norms' Trust Model

This concept is validated by proposing a norm's trust model based on an agent's belief about Authority, Reputation, and Adoption for adopting the norms in a new environment.

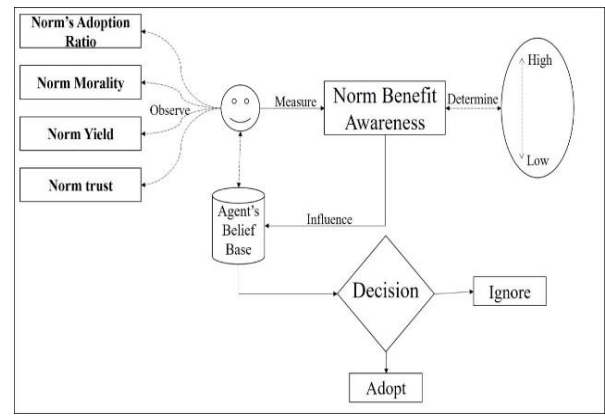


Fig. 1. Evaluating the Norm's benefit awareness.

B. Authority

A factor that determines the trust value of a particular norm is observing authorized agents, which is one of the resources for a new agent when joining a society. Authorized agents represent their societies and have the authority to reward or sanction a society's member. Therefore, authorized agents are trusted and its norm has a high trust value. The verification is justified by an agent, which endorses the norm indicating that the norm is trusted by the authorized body.

Therborn [18] states that the acceptance of a particular norm is significantly greater if the individual views the source as being credible, such as an accredited/prestigious organization, parents or people in authority. However, we exploit the agent's authority level proposed by Abdul Hamid et al. [19], who divide the trust level into three categories; low, medium and high. While Abdul Hamid et al. [19] divide the trust level into three categories, we exploit only two categories: Trust (1) and Distrust (0).

Definition 2: Authority, \mathcal{A}_A , is a set of agents in the Domain D , which have the power that derives its legitimacy by respecting cultural patterns and existing rules and regulations, such as Governments.

If \mathcal{A}_A is a set of Authority agents, and \mathbb{A}_A is an Authority agent, then,

$$\mathcal{A}_A = (\mathbb{A}_{A1}, \mathbb{A}_{A2}, \dots, \mathbb{A}_{An})$$

If \mathbb{A}_A is an authorized agent in the domain, D , then

$$\mathbb{A}_A \in \mathcal{A}_A \Leftrightarrow \text{Authorised}(\mathbb{A}_A, D) \wedge \mathbb{A}_A \in D \quad (1)$$

This means that \mathbb{A}_A belongs to authorized agents \mathcal{A}_A , if and only if, \mathbb{A}_A is authorized in D and \mathbb{A}_A belongs to D . In this regard, a visitor agent asks the authorized agents about a candidate norm, \hat{C}_{η} whether or not it is trusted and determines the summation of authorised agents \mathcal{A}_A , and if the summation is ONE then return value ONE (1), otherwise, return ZERO (0).

$$\mathcal{A}_A(\hat{C}_{\eta}) = \begin{cases} \hat{C}_{\eta} = 1 & \text{if } \sum_{i=j}^j \mathbb{A}_A(j) = 1 \\ \hat{C}_{\eta} = 0 & \text{otherwise} \end{cases} \quad (2)$$

C. Reputation

Reputation is not an expectation without bounds but learning of the past. A sociologist, Barbara Misztal [20], states that reputation is a memory fixed to a particular personality. Simply, a strong reputation builds trust and thus a type of social evaluation. It is a conviction about other's assessment. Josang et al. [21] describe reputation as an opinion about an entity, therefore, interactions between people generate reputation. Experience gained from interactions between members of a society sets reputation values for others.

Shinji [22] shows that agents will be motivated due to reputation formation. Abdul Hamid et al. [19] believe that the reputation of an agent, which practices a norm in a new environment, impacts the norm's trust value. The Neighbour-Trust Algorithm is exploited to calculate the reputation score of each agent [14].

$$r_{VN} = \frac{\sum_{i \in \text{neighbor}(N)} w_{Vi} \cdot t}{\sum_{i \in \text{neighbor}(N)} w_{Vi}} \quad (3)$$

where r_{VN} is the reputation value, t is the direct trust values of N neighbouring agents, and w_{Vi} are the weights that represent the personal opinion of the requesting agent. These weights are normally independent of the context of the direct trust values the neighbors provide.

For example, if a visitor agent, A, wants to get information about agent C, agent A asks agent B about its opinion on agent C. In this case, w_{Vi} is the trust weight that agent A gives based on the information which agent B provides. t is the direct trust value agent B has about agent C. Later, when agent A might have a direct experience with agent C, the trust value is represented by t only. To get a more accurate value, agent A should ask many more neighbour agents.

D. Adoption Ratio

A Norm Adoption Ratio (N_{AR}) is the ratio of agents practicing a particular norm to the population of agents in a community. To calculate the N_{AR} , a formula proposed by Mahmoud et al. [10] is used. The formula is called a Norm Strength (NS). In their work, they assume that an agent observes a society's members' activities, collects episodes and add these to a record file to be analyzed for detecting the potential norms. The episode is a set of events that an agent enacts in a domain to achieve its goal. For example, in a restaurant domain, the episode might be "arrive, sit, order, eat, pay, tip, and depart" [23].

The calculation of the Norm Strength according to Mahmoud et al. [10], is as follows, where n is a norm:

$$NS(n) = \frac{\text{Number of episodes which include } n}{\text{Total number of episodes}} \quad (4)$$

From Fig. 2, there is an agent and a number of norms. The agent first (1) observes the norms of an environment. Then, it (2) detects the potential norm and (3) evaluates the norm based on Authority, Reputation, and Adoption to obtain the norm's trust value. The agent then (4) updates the norm's trust value of the detected norms to its belief base (5). The agent can reason and decide to comply with or even adopt the potential norm.

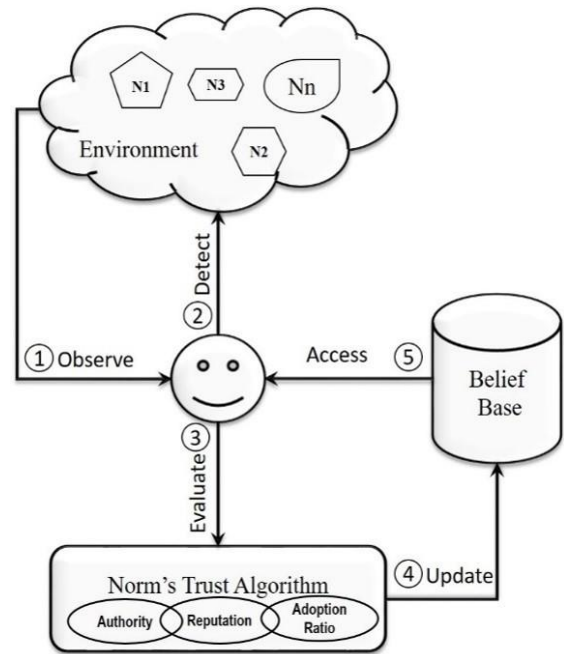


Fig. 2. The Norm's Trust Model.

The norm's trust algorithm assesses the Authority, Reputation and Adoption Ratio of the potential norm to evaluate the norms' trust value. The norms' trust value contributes to the adopt/reject decision.

VI. NORM'S TRUST EVALUATION MODEL

Abdul Hamid et al. [19] propose a norm's trust concept, which is based on the transitive trust of a visitor agent who trusts a local agent's information of another local agent enacting a detected norm. This concept is exploited using the three factors associated with the process: Authority, Reputation, and Adoption Ratio.

Fig. 3 illustrates the trust inference process that applies to a particular norm. Agent A firstly observes a set of behaviours which agents B, C and D perform. Then, agent A infers the norm's trust value of the norm, $n1$, if agents B, C, and D perform the norm, $n1$. Through the three filters that influence the norm's trust, agent A evaluates the trustworthiness of the agents B, C and D and infers the norm $n1$'s trust value.

Based on literature relating to trust and reputation models of MAS [17], [24], a number of information sources, namely anecdotal evidence, personal/direct experience, witness accounts, and social studies data that play a role in affecting a trust value are ascertained. For the purpose of evaluating the trustworthiness of agents, the trust factor as defined in the context of these models, have been used. In this research, the motives for adopting the norms, together with analysis of the mentioned sources, are both given due importance.

Based on these analyses, three main factors are categorized that influence norms existence in a society, which are Authority, Reputation, and Adoption Ratio that are mentioned earlier.

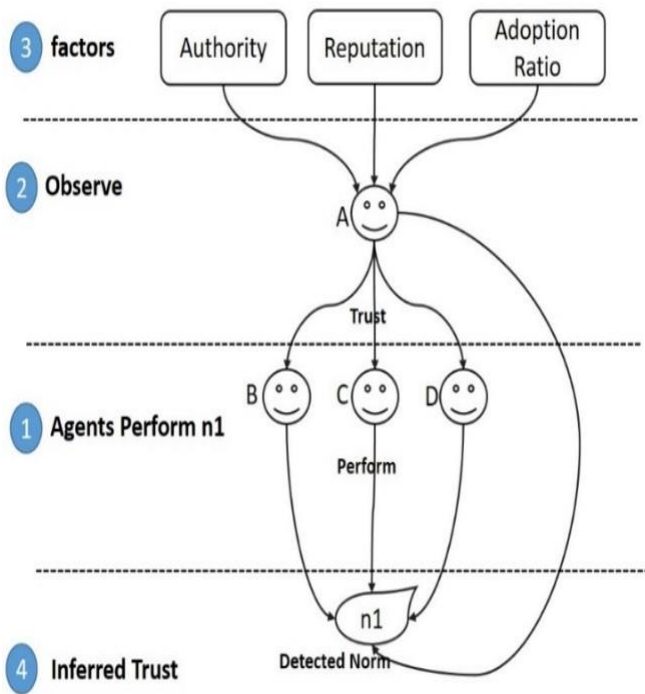


Fig. 3. Trust Inferences through Filters.

A trust value influences the decisions which can be determined from the identified factors. To determine the norms' trust (NT) value, we consider the three factors (Authority, A; Reputation, R; and Adoption Ratio, AR). We assume that the threshold value for a norm trust value, $NT = 0.5$. While Abdul Hamid et al. [19] describe three levels of a norm's trust, in this work only two levels are exploited:

- **Trust, NT_F :** A norm is fully trusted when all the three parameters (A, R, AR) each holds a value that jointly produces a high value of the norm's trust. There is no conflict between the values of the parameters and the agent positively verifies the norm with all factors. An agent, α , entirely trusts the norm, η , if and only if all the three parameters indicate high values of trust in the norm, η :

$$NT_F(\alpha, \eta) \leftrightarrow \text{trust}((A \wedge R \wedge AR), \eta) \quad (5)$$

- **Distrust, NT_D :** A norm is distrusted when all the three parameters negatively produce a very low value. This means that the agent, α , distrusts the norm, η , if and only if all the three parameters indicate low values of trust in the norm, η .

$$NT_D(\alpha, \eta) \leftrightarrow \text{distrust}((A \wedge R \wedge AR), \eta) \quad (6)$$

Therefore, the formulation of a decision to Trust or Distrust is as follows:

For an agent, α , the detected norm, η , a Trust decision is 1, and Distrust decision is 0:

$$NT(\eta\alpha) = \begin{cases} 1, & NTF \geq 0.5 \\ 0, & NTD < 0.5 \end{cases} \quad (7)$$

TABLE I. THE SUMMARY OF NORM ADOPT/REJECT DECISION

Condition	Norm's Trust (NT) Level	
		Decision
$NT = 1$	Trust	The agent <i>will</i> adopt a norm if its norm's trust value is equal to the highest possible value, 1. $NT_F: NT = 1$
$NT = 0$	Distrust	An agent <i>will</i> reject a norm if its norm's trust value is equal to the lowest possible value, 0. $NT_D: NT = 0$

These decisions are shown as a willingness matrix that portrays the adoption or rejection of a norm. The willingness level to adopt or reject depends on the NT threshold value (0.5). Table I shows the summary of the decision's options.

VII. SOCIAL SIMULATION

An example of a social simulation is presented, in which a visitor agent, A, enters a train station to take a train to another station. Agent A observes other local agents' behaviours in the domain and through its norm detection function, agent A detects three different behaviours practiced by the local agents which are; 11 agents queue and wait behind a yellow line (N_1), five agents wait while sitting on a bench (N_2), and four agents loiter around the platform (N_3). Agent A has to decide which behaviour it has to trust and adopt.

In this example, the first stage in a norm's trust evaluation, agent A evaluates its neighbours' norm trust values based on the reputation scores using (3) and the authority level [18]. Based on the Neighbour-Trust Algorithm [14] to calculate the trust level for norm n1, agent A evaluates the reputation score for Agent1 at this stage, by asking the neighbour agents' opinions about Agent1. It is assumed that the visitor agent A obtains all the reputation values, t . It then assigns the corresponding weights, w_{Vi} as shown in Table II below for each of the neighbor agents. Based on (3) the visitor agent calculates the reputation score of the potential norm.

TABLE II. REPUTATION SCORE OF NEIGHBOUR AGENTS

Agent1 Neighbor's	w_{Vi}	t	$w_{Vi} \cdot t$
Agent2	0.99	0.92	0.9108
Agent4	0.88	0.80	0.7040
Agent8	0.89	0.88	0.7832
Agent11	0.77	0.90	0.6930
Agent15	0.66	0.88	0.5808
Agent16	0.75	0.88	0.6600
Agent19	0.95	0.88	0.8360
Sum	5.89	6.14	5.1678
	r_{VN}		0.87739

From the table, the reputation score of Agent1 is 0.87739, which is a high reputation.

In the second stage, agent A evaluates the authority level of Agent1 based on agent A’s database. Consequently, the Authority is (1). Then, in the third stage, agent A evaluates the Adoption Ratio. As mentioned earlier, the trust value of the potential norms (NT) is calculated based on its Adoption Ratio, AR. Using (4), the list of Reputation Scores and Authority for each neighbour and the Adoption Ratio for each potential norm is as listed in Table III. This shows the values of Reputation, Authority and the adoption Ratio of each potential norm practiced by the neighbour agents. Consequently, the visitor agent decides to adopt the norm, n1, as it is the only trusted behavior.

The trust model is validated as a simulation of the train station scenario by using Netlogo, which is a programmable agent-based modelling environment for simulating natural and social phenomena. The simulation is run five times and each run has a new environment with a different number of norms (see Fig. 4). In each run, the visitor agent observes and detects the norms in the environment, calculates and evaluates the trust value for the potential norm and decides whether to trust or distrust it.

Based on these premises, Table IV shows the simulation results. The results show that in Runs 1 and 3, the trusted norm is SIT, while in Runs 2 and 4, QUEUE is the trusted norm. Hence a visitor agent may adopt these two norms in this particular environment.

TABLE III. THE TRUST VALUE OF POTENTIAL NORMS

Practicing Agents	Norm, n_i	Neighbor, N_i	Reputation Score	Authority Level	Adoption Ratio, AR	Trust Level
Agent1	n1	N1	0.87	1	0.55	Trusted
Agent2	n1	N2	0.45	0	0.55	Distrust
Agent3	n1	N3	0.4	0	0.55	Distrust
Agent4	n1	N4	0.43	0	0.55	Distrust
Agent5	n1	N5	0.49	0	0.55	Distrust
Agent6	n1	N6	0.43	0	0.55	Distrust
Agent7	n1	N7	0.49	0	0.55	Distrust
Agent8	n1	N8	0.45	0	0.55	Distrust
Agent9	n1	N9	0.81	1	0.55	Trusted
Agent10	n1	N10	0.43	0	0.55	Distrust
Agent11	n1	N11	0.39	0	0.55	Distrust
Agent12	n2	N12	0.36	0	0.41	Distrust
Agent13	n2	N13	0.33	0	0.41	Distrust
Agent14	n2	N14	0.38	0	0.41	Distrust
Agent15	n2	N15	0.31	0	0.41	Distrust
Agent16	n2	N16	0.44	0	0.41	Distrust
Agent17	n3	N17	0.49	0	0.33	Distrust
Agent18	n3	N18	0.45	0	0.33	Distrust
Agent19	n3	N19	0.42	0	0.33	Distrust
Agent20	n3	N20	0.23	0	0.33	Distrust

TABLE IV. SIMULATION RESULTS

Simulation Runs	Potential Norm	Adoption Ratio	Authority	Reputation	Trust Value	Decision
Run 1	SIT	1	1	1	1	Trust
	QUEUE	0	1	0	0	Distrust
	LOITER	1	0	1	0	Distrust
Run 2	SIT	0	0	0	0	Distrust
	QUEUE	1	1	1	1	Trust
	LOITER	0	0	1	0	Distrust
Run 3	SIT	1	1	1	1	Trust
	QUEUE	0	1	1	0	Distrust
	LOITER	0	0	0	0	Distrust
Run 4	SIT	0	0	1	0	Distrust
	QUEUE	1	1	1	1	Trust
	LOITER	0	1	1	0	Distrust
Run 5	SIT	0	1	1	0	Distrust
	QUEUE	0	0	0	0	Distrust
	LOITER	0	0	1	0	Distrust

The findings in this research are significant in that they offer an elaborate approach to norms’ analysis and computation for an eventual norm’s adoption or rejection in normative multi-agent systems. The norm’s adoption or rejection is based on the computation of the norms’ factors which manifest the benefits that the norms would entail to achieve the agents’ goals. Consequently, these findings significantly contribute to the literature in normative multi-agent systems.

VIII. SIMULATION MODEL

In this simulation model, using NetLogo designed by Uri Wilensky (1999), a virtual environment is created for calculating the norm’s trust. The virtual environment is a train station which is represented by the passengers (people) and the inspector (visitor agent). The virtual environment has the functions to create a new domain, select and run a domain, and set the variables of the domain. Fig. 4 shows the user interface after opening and running a model from the Models Library. It has three parts which are:

- The top left part of the window shows the train station environment, which consists of passenger agents, senior agents, authorized agents and the visitor agent.
- The left part of the window (text box) shows the results of the norms that the agent detected. The text box shows the values of all the potential norms. It also shows the procedure of the Norm’s Trust calculation for the potential norms in the domain.
- The bottom left part of the window shows the simulation buttons for controlling the simulation model and has a few boxes and buttons which are:

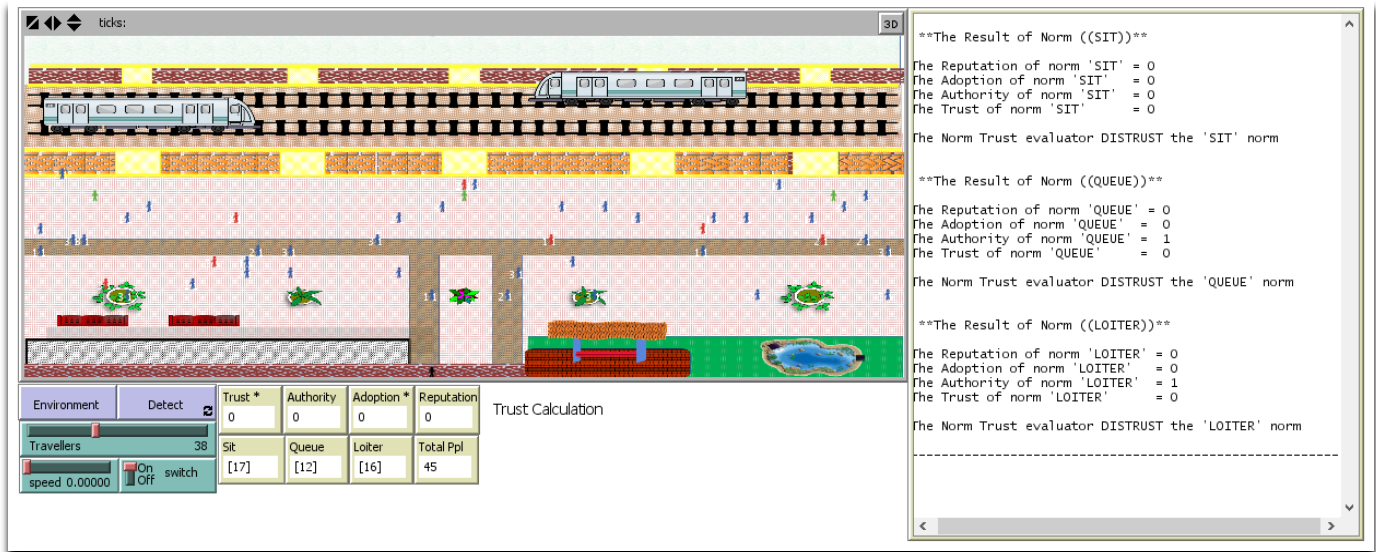


Fig. 4. Trust modelling simulation.

- **Environment:** To create a new instance of the train station. By clicking the button, a new environment is configured. The number of passengers is selected from a slider, labeled as Traveler. A number of senior agents is created randomly. Three policemen are deployed on the station platform. In addition, a visitor agent is also created. Every time the button is clicked, it creates a new instance of the domain.
- **Detect:** To enable the visitor agent to detect the potential norms and calculate the norms' trust values. When the button is clicked, the visitor agent tours the station platform and collects as much information as possible about the enacted norms by the passengers and return to its original location to analyze the collected data.
- **Travelers:** To set the number of passengers (people) at the station. By using this slider the number of passengers on the platform from 0 up to 100 passengers is created.
- **Speed:** To set the speed of the simulation using the slider.
- **Trust:** Shows the result of a trusted norm of the potential norms based on the values of Authority, Adoption Ratio, and Reputation of the potential norms. Thus, the trust is either "1" which means the potential norm is a Trusted norm or "0" which means the potential norm is a Distrusted norm.
- **Authority:** Shows the norm's trusted value of the potential norm according to authorized agents' opinion. Based on that recommendation, the result is either "1" which means the norm is trusted or "0" which means the norm is distrusted.
- **Adoption:** Shows the result of adoption ratio of the potential norm based on the number of people who trust and practice the potential norm. So, if the number of agents who trusts the potential norm is more than 50% then the adoption result shows "1"

- or "0" if the number of agents who trust the potential norm is less than 50%.
- **Reputation:** Shows the norm's reputation value of the potential norm according to senior agents' belief. Based on that, it shows "1" if the potential norm has a high reputation or "0" if it has a low reputation.

IX. CONCLUSIONS AND FURTHER WORK

In this paper, a norm's trust model is proposed to facilitate agents' decision-making process in norm adoption or rejection. The model constitutes a technique that assists agents in determining the norm's benefits to improve agents' decisions in adopting or rejecting the norms. A norm's trust formula is exploited based on Abdul Hamid et al. [19], but a new architecture is proposed for calculating the norm's trust. The model is validated by a simulation, in which a visitor agent observes other local agents' behaviours in a train station and detects three different behaviours enacted by the local agents. The simulation results indicate that the trust model imparts a trustable value for the detected norms, which the agent can use to adopt or reject the norms.

In cases where agents encounter multiple norms, the norms' trust levels indicate how much they can be relied upon in fulfilling the normative goals (generated from the adopted norms), neither conflicting with the agents' internal structures nor interfering with their intended goals.

This paper is a part of the authors' research in agent 'awareness' of norms' benefits. A norm's trust is an important factor, whose value is needed to be determined as a parameter in the formulation of a norm's benefits. The benefits are a measure with which a decision is made whether to adopt or reject a detected norm. The other parameters in an earlier publication [8] are Norm's Adoption Ratio, Norm's Yield, and Norm's Morality.

In future works, all these parameters shall be included in the formulation of the norm's benefits and a comprehensive simulation to validate the formulation shall be developed.

REFERENCES

- [1] S. Parsons, K. Atkinson, K. Haigh, K. Levitt, P. M. J. Rowed, M. P. Singh, et al., "Argument schemes for reasoning about trust," *Computational Models of Argument: Proceedings of COMMA 2012*, vol. 245, p. 430, 2012.
- [2] D. H. McKnight and N. L. Chervany, "The meanings of trust," 1996.
- [3] J.-M. Seigneur and P. Dondio, "Trust and reputation for successful software self-organisation," in *Self-organising Software*, ed: Springer, 2011, pp. 163-192.
- [4] R. Falcone and C. Castelfranchi, "Social trust: A cognitive approach," in *Trust and deception in virtual societies*, ed: Springer, 2001, pp. 55-90.
- [5] D. M. Romano, "The nature of trust: conceptual and operational clarification," Louisiana State University, 2003.
- [6] S. C. Curral and M. J. Epstein, "The Fragility of Organizational Trust:: Lessons From the Rise and Fall of Enron," *Organizational Dynamics*, vol. 32, pp. 193-206, 2003.
- [7] N. H. A. Hamid, M. S. Ahmad, A. Ahmad, A. Mustapha, M. A. Mahmoud, and M. Z. M. Yusoff, "Trusting Norms: A Conceptual Norms' Trust Framework for Norms Adoption in Open Normative Multi-agent Systems," in *Distributed Computing and Artificial Intelligence*, 12th International Conference, S. Omatu, Q. M. Malluhi, S. R. Gonzalez, G. Bocewicz, E. Bucciarelli, G. Giulioni, et al., Eds., ed Cham: Springer International Publishing, 2015, pp. 149-157.
- [8] A.-M. K. Itaiwi, M. S. Ahmad, M. A. Mahmoud, and A. Y. Tang, "Norm's Benefit Awareness in Open Normative Multi-agent Communities: A Conceptual Framework," in *Distributed Computing and Artificial Intelligence*, 11th International Conference, 2014, pp. 209-217.
- [9] A. Artikis, M. Sergot, and J. Pitt, "Specifying norm-governed computational societies," *ACM Transactions on Computational Logic (TOCL)*, vol. 10, p. 1, 2009.
- [10] M. A. Mahmoud, A. Mustapha, M. S. Ahmad, A. Ahmad, M. Z. M. Yusoff, and N. H. A. Hamid, "Potential norms detection in social agent societies," in *Distributed Computing and Artificial Intelligence*, ed: Springer, 2013, pp. 419-428.
- [11] B. T. R. Savarimuthu, S. Cranefield, M. A. Purvis, and M. K. Purvis, "Identifying prohibition norms in agent societies," *Artificial intelligence and law*, vol. 21, pp. 1-46, 2013.
- [12] J. Duffy, H. Xie, and Y.-J. Lee, "Social norms, information, and trust among strangers: theory and evidence," *Economic theory*, pp. 1-40, 2013.
- [13] M. Van Dijke, D. De Cremer, and D. M. Mayer, "The role of authority power in explaining procedural fairness effects," *Journal of Applied Psychology*, vol. 95, p. 488, 2010.
- [14] R. Kiefhaber, S. Hammer, B. Sava, J. Schmitt, M. Roth, F. Kluge, et al., "The neighbor-trust metric to measure reputation in organic computing systems," in *Self-Adaptive and Self-Organizing Systems Workshops (SASOW)*, 2011 Fifth IEEE Conference on, 2011, pp. 41-46.
- [15] R. Falcone, C. Castelfranchi, H. L. Cardoso, A. Jones, and E. Oliveira, "Norms and Trust," in *Agreement Technologies*, ed: Springer, 2013, pp. 221-231.
- [16] P. Faulkner, "Norms of trust," 2010.
- [17] G. Andrighetto, D. Villatoro, and R. Conte, "Norm internalization in artificial societies," *Ai Communications*, vol. 23, pp. 325-339, 2010.
- [18] G. Therborn, "Back to norms! On the scope and dynamics of norms and normative action," *Current Sociology*, vol. 50, pp. 863-880, 2002.
- [19] N. H. A. Hamid, M. S. Ahmad, A. Ahmad, A. Mustapha, M. A. Mahmoud, and M. Z. M. Yusoff, "Trusting Norms: A Conceptual Norms' Trust Framework for Norms Adoption in Open Normative Multi-agent Systems," in *Distributed Computing and Artificial Intelligence*, 12th International Conference, 2015, pp. 149-157.
- [20] B. Misztal, *Trust in modern societies: The search for the bases of social order*: John Wiley & Sons, 2013.
- [21] A. Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," *Decision support systems*, vol. 43, pp. 618-644, 2007.
- [22] S. Teraji, "A theory of norm compliance: Punishment and reputation," *The Journal of Socio-Economics*, vol. 44, pp. 1-6, 2013.
- [23] B. T. R. Savarimuthu, "Mechanisms for norm emergence and norm identification in multi-agent societies," University of Otago, 2011.
- [24] I. Pinyol and J. Sabater-Mir, "Computational trust and reputation models for open multi-agent systems: a review," *Artificial Intelligence Review*, vol. 40, pp. 1-25, 2013.

AUTHORS' PROFILES

Al-Mutazbellah K. Itaiwi has obtained his B.Sc. in Computer Science from the College of Computer, University of Anbar, Iraq in 2007. he obtained his Master of Information Technology at the College of Graduate Studies, Universiti Tenaga Nasional (UNITEN), Malaysia in 2012 and enrolled in the PhD of Information and Communication Technology program since 2013 at the College of Graduate Studies, Universiti Tenaga Nasional (UNITEN), Malaysia. During his studentship at UNITEN, he conducted additional laboratory work for degree students at the College of Engineering. his current research interests include software agents and multi-agent systems

Mohd Sharifuddin Ahmad is currently the Head of Center for Agent Technology (CAT) at the College of Computer Science and Information Technology, Universiti Tenaga Nasional (UNITEN). He obtained his MSc. in Artificial Intelligence from Cranfield University, UK in 1995. He obtained his PhD. in Artificial Intelligence from Imperial College, London, UK in 2005. His research interests include Software Agents and Knowledge Management.

Alicia Y.C. Tang currently works at the Department of Systems and Networking, Universiti Tenaga Nasional (UNITEN). Alicia does research in Agents & Autonomous Systems, Data Mining and Artificial Intelligence. Their current project is 'i-VSM and i-RAM based on the concept of Agents of Things (AoT).'