The utility of categories through their recommendation in an agents world with local or not local communication

Rino Falcone (rino.falcone@istc.cnr.it)
Institute for cognitive Science and Technology, ISTC-CNR, Rome

Alessandro Sapienza (alessandro.sapienza@istc.cnr.it)
Institute for cognitive Science and Technology, ISTC-CNR, Rome

Cristiano Castelfranchi (cristiano.castelfranchi@istc.cnr.it)
Institute for cognitive Science and Technology, ISTC-CNR, Rome

Abstract

In this paper we focus on the importance of generalized knowledge: agents' categories. The cognitive advantage of generalized knowledge can be synthesized in this claim: "It allows us to know a lot about something/somebody we do not directly know". At a social level this means that I can know a lot of things on people that I never met; it is social "prejudice" with its good side and fundamental contribution to social exchange.

In this study we experimentally inquire the role played by categories' reputation with respect to the reputation and opinion on single agents: when it is better to rely on the first ones and when are more reliable the second ones.

We will consider two different scenarios: one strongly influenced by the spatial distance between agents (localized world); the other totally independent by the spatial distances (non-localized world), quite similar to the modern web society, in which the communicative distance follows different routs with respect to the spatial distance.

We want to investigate how the parameters defining the specific environment (number of agents, their interactions, transfer of reputation, and so on) influence the importance of categories' reputation in these two different worlds.

Keywords: trust; cognitive analysis; social simulations.

Introduction

Knowing without knowing

Knowledge generalization and its organization around "classes" of entities and events (Falcone and Castelfranchi, 2008) is a foundational need of human cognition.

According to us a category is a homogeneous group of things (agents in this case) identified or inferred by a set of visible and non-misleading *signs*. This is why coats, uniforms, titles, badges, diplomas, etc. are so important in social life and it is crucial their exhibition and the assurance of their authenticity (and, on the other side, the ability to falsify and deceive). Examples of categories are: dogs, cat, doctors, sellers, thieves etc.

In this work we assume that the membership to a given class or category is true and transparent: the category of a given agent is public, common knowledge. We are just interested in the fact that an agent belongs to a category.

The advantage of such a *hierarchic structure of knowledge* is not only economical: we do not reproduce beside and for each "instance" in our memory. We just write that around the category and then - when needed - instantiate it on a specific object.

The greatest advantage is not just in memory space and costs, but in the fact that we *know a lot of thing about something that we never met*; just by inference, prediction, inheritance. We have a lot of knowledge about a given entity without any direct experience on it. This crucial power of our cognitive organization is obviously exploited also in social life, in order to have information and expectations about people that we never met.

This fundamental device for "knowing without knowing" is surely crucial also for *trust evaluations*. Society works also on the basis of trust between strangers; this trust is based on several inferential and social tricks (like evoked feelings, analogy, recommendations, etc.) but is also strongly relying on *categories* of people and their "signaling" and recognition. If we (dis-)trust a given class of people and we understand that Y belongs to that class we can (dis-)trust Y.

The problems about categories are:

- How do we build our trust in a category? From our direct experience or trust in its members? How many of them are necessary in order to generalize?
- How much risky is the instantiation from the class to that member Y? How much reliable are "signals" about Y membership? How much Y is representative, typical, of that class? And how much variance of trustworthiness there is in that class?
- When and how much it is advantageous to exploit trust on the categories and not just direct trust in the individual?

In this study we intend to explain and experimentally show the advantage of trust evaluation based on classes' reputation with respect to the reputation and opinion on single potential agents (partners). In an open world or in a broad population how can we have sufficient direct or reported experience on everybody? The quantity of potential agents in that population or net that might be excellent partners but that nobody knows enough can be high.

Our claim is that: the larger the population and the ignorance about the trustworthiness of each individual the more precious the role of trust in categories. If I know (through signals, marks, declaration, ...) the class of a given guy/agent I can have a reliable opinion of its trustworthiness derived from its class-membership.

It is clear that the advantages of such cognitive power provided by categories and prejudices do not only depend on recommendation and reputation about categories. We can personally build - by generalization - our evaluation of a given category from our direct experience with its members (this fact happens in our experiments for the agents that later have to propagate their recommendation about). However, in this simulation we have in the trustor (which has to decide whom rely on) only a prejudice based on recommendations about that category and not its personal experience.

Under a certain degree on direct experiences and circulation of recommendations, the performance of the evaluation based on classes will perform better; and in certain cases there will be no alternative at all: we do not have any evaluation on that individual, a part from its category; either we work on inferential instantiation of trustworthiness or we loose a lot of potential partners. This powerful inferential device has to be strongly present in WEB societies supported by MAS. We simplify here the problem of the generalization process, of how to form judgment about groups, classes, etc. by putting aside for example inference from other classes (higher or sub); we build opinion (and then its transmission) about classes on the bases of experience with a number of subjects of a given class.

In this work we are also interested in showing the difference between localized and non-localized knowledge. A localized world is a world strongly influenced by the spatial distance between agents; a non-localized world is independent by the spatial distances, in which the communicative distance follows different routs with respect to the spatial distance. The first approach reflects the traditional social way to exchange information, before the advent of virtual communities, where communication is constrained by spatial distance.

However, nowadays we also use another way to exchange information: the Web. Here we have access to a more complex net of users; our choice follows (and is influenced) by different communicative links to the information sources.

We are interested in analyzing the utility of categories in this two different contexts, trying to understand if and how they affect its performance.

Related works

Differently from (Burnett et al, 2010; Fang et al, 2012; Sensoy et al, 2014), in this work we do not address the problem of learning categorical knowledge and we assume that the categorization process is objective. Similarly to (Burnett et al, 2013), we give agents the possibility to recommend categories.

In the majority of the cases available in the literature, the concept of recommendation is used concerning recommender systems (Adomavicious et al, 2015). These ones can be realized using both past experience (content-based RS) (Lops et al, 2011) or collaborative filtering, in which the contribute of single agents/users is used to provide group recommendations to other agents/users.

A classical decentralized approach is referral systems (Yolum and Singh, 2003), where agents adaptively give referrals to one another.

Information sources come into play in FIRE (Huynh et al, 2006), a trust and reputation model that use them to produce a comprehensive assessment of an agent's likely performance.

The described solutions are quite similar to our work, although we contextualized this problem to information sources. However we do not investigate recommendations with just the aim of suggesting a particular trustee, but also for inquiring categories' recommendations.

Recommendation and reputation: definitions

Let us consider a set of agents $Ag_1, ..., Ag_n$ in a given world. We consider that each agent in this world could have trust relationships with anyone else. On the basis of these interactions the agents can evaluate the trust degree of their partners, so building their judgments about the trustworthiness of the agents with whom they interacted in the past.

The possibility to access to these judgements, through recommendations, is one of the main sources for trusting agents outside the circle of closer friends. Exactly for this reason recommendation and reputation are the more studied and diffused tools in the trust domain (Ramchurn et al, 2004) .

We introduce

$$\mathrm{Rec}_{x,y,z}(\tau) \tag{1}$$

where $x, y, z \in \{Ag_1, Ag_2, ..., Ag_n\}$, we call D the specific set of agents: $D \equiv \{Ag_1, Ag_2, ..., Ag_n\}$

and $0 \le \mathrm{Rec}_{x,y,z}(\tau) \le 1$

τ, as established in the trust model of (Castelfranchi and Falcone, 2010), is the task on which the recommender x expresses the evaluation about y.

In words: $\mathrm{Rec}_{x,y,z}(\tau)$ is the value of x's recommendation about y performing the task τ, where z is the agent receiving this recommendation. In this paper, for sake of simplicity, we do not introduce any correlation/influence between the value of the recommendations and the kind of the agent receiving it: the value of the recommendation does not depend from the agent to whom it is communicated.

So (1) represents the basic expression for recommendation.

We can also define a more complex expression of recommendation, a sort of *average recommendation*:

$$\sum_{x=Ag_1}^{Ag_n} \mathrm{Rec}_{x,y,z}(\tau)/n \tag{2}$$

in which all the agents in the defined set of agents express their individual recommendation on the agent y with respect the task τ and the total value is divided by the number of agents.

We consider the expression (2) as the <u>reputation</u> of the agent y with respect to the task τ in the set D.

Of course the reputation concept is more complex than the simplified version here introduced (Conte and Paolucci, 2002; Sabater-Mir, 2003).

It is in fact the value that would emerge in the case in which we receive from each agent in the world its recommendation about y (considering each agent as equally reliable).

In the case in which an agent has to be recommended not only on one task but on a set of tasks (τ_1 , ..., τ_k), we could define instead of (1) and (2) the following expressions:

$$\sum_{i=1}^{k} \mathrm{Re}c_{x,y,z}(\tau_i)/k \quad (3)$$

that represents the x's recommendation about y performing the set of tasks ($\tau_1,..., \tau_k$), where z is the agent receiving this recommendation.

Imagine having to assign a meta-task (composed of a set of tasks) to just one of several agents. In this case the information given from the formula (3) could be useful for selecting (given the x's point of view) on average (with respect to the tasks) the more performative agent y.

$$\sum_{x=Ag_1}^{Agn} \sum_{i=1}^{k} \mathrm{Re}c_{x,y,z}(\tau_i)/nk \quad (4)$$

that represents a sort *of average recommendation* from the set of agents in D, about y performing the set of tasks (τ_1 , ..., τ_k). *We consider the expression (4) as the reputation of the agent y with respect the set of tasks (τ_1 , ..., τ_k), in the set D.*

Having to assign the meta-task proposed above, the information given from the formula (4) could be useful for selecting on average (with respect to both the tasks and the agents) the more performative agent y.

Using Categories

As described above, an interesting approach for evaluating agents is to classify them in specific categories already pre-judged/rated and as a consequence to do inherit to the agents the properties of their own categories.

So we can introduce also the *recommendations about categories*, not just about agents (we discuss elsewhere how these recommendations are formed). In this sense we define:

$$\mathrm{Re}c_{x,Cy,z}(\tau) \quad (5)$$

where $x \in \{Ag_1, Ag_2, ..., Ag_n\}$ as usual, and we characterize the categories $\{C_1, ..., C_l\}$ through a set of features $\{f_{y1}, ..., f_{ym}\}$:

$$\forall y \in \{Ag_1, ..., Ag_n\} \exists c_y \in \{C_1, ..., C_l\} \mid (C_y \equiv \{f_{y1}, ..., f_{ym}\}) \wedge (\{f_{y1}, ..., f_{ym}\} \in y)$$

it is clear that there is a relationship between task τ, and the features $\{f_{y1}, ..., f_{ym}\}$ of the C_y category. In words we can say that each agent in D is classified in one of the categories $\{C_1, ..., C_l\}$ that are characterized from a set of features $\{f_1, ..., f_m\}$; as a consequence each agent belonging to a category owns the features of that category. $0 \le \mathrm{Re}c_{x,Cy,z}(\tau) \le 1$

In words: $\mathrm{Re}c_{x,Cy,z}(\tau)$ *is the value of x's recommendation about the agents included in category C_y when they perform the task τ*, (as usual z is the agent receiving this recommendation).

We again define a more complex expression of recommendation, a sort of *average recommendation*:

$$\sum_{x=Ag_1}^{Ag_n} \mathrm{Re}c_{x,Cy,z}(\tau)/n \quad (6)$$

in which all the agents in the domain express their individual recommendation on the category C_y with respect the task τ and the total value is divided by the number of the recommenders.

We consider the expression (6) as the reputation of the category C_y with respect the task τ in the set D.

Now we extend to the categories, in particular to C_y, the recommendations on a set of tasks (τ_1, ..., τ_k):

$$\sum_{i=1}^{k} \mathrm{Re}c_{x,Cy,z}(\tau_i)/k \quad (7)$$

that represents *the recommendation value of the x's agent about the agents belonging to the category C_y when they perform the set of tasks (τ_1,...,τ_k).*

Finally, we define:

$$\sum_{x=Ag_1}^{Agn} \sum_{i=1}^{k} \mathrm{Re}c_{x,Cy,z}(\tau_i)/nk \quad (8)$$

that represents *the value of the reputation of the category C_y (of all the agents y included in C_y) with respect the set of tasks (τ_1,...,τ_k), in the set D.*

Definition of Interest for this Work

In this paper we are in particular interested in the case in which z (a new agent introduced in the world) asks for recommendation to x ($x \in D$) about an agent belonging to its domain D_x for performing the task τ (D_x is a subset of D, it is composed by the agents that x knows). x will select the best evaluated y, with $y \in D_x$ on the basis of formula:

$$\max_{y \in D_x}(\mathrm{Re}c_{x,y,z}(\tau)) \quad (9)$$

where $D_x \equiv \{Ag_1, Ag_2, ..., Ag_m\}$, D_x includes all the agents evaluated by x. They are a subset of D: $D_x \subseteq D$.

In general D and D_x are different because x does not necessarily know (has interacted with) all the agents in D.

z asks for recommendations not only to one agent, but to a set of different agents: $x \in D_z$ (D_z is a subset of D, to which z asks for reputation), and selects the best one on the basis of the value given from the formula:

$$\max_{x \in D_z}(\max_{y \in D_x}(\mathrm{Re}c_{x,y,z}(\tau))) \quad (10)$$

$D_z \subseteq D$, z could ask to all the agents in the world or to a defined subset of it (see later).

We are also interested to the case in which z ask for recommendations to x about a specific *agents' category* for performing the task τ. x has to select the best evaluated C_y among the different $C_y \in \{C_1, ..., C_l\}$ x has interacted with (we are supposing that each agent in the world D, belongs to a category in the set $\{C_1, ..., C_l\}$).

In this case we have the following formulas:

$$\max_{Cy \in D_x}(\mathrm{Re}c_{x,Cy,z}(\tau)) \quad (11)$$

that returns the category best evaluated from the point of view of an agent (x). And

$$\max_{x \in D_z}(\max_{Cy \in D_x}(\mathrm{Re}c_{x,Cy,z}(\tau))) \quad (12)$$

that returns the category best evaluated from the point of view of all the agents included in D_z.

Computational Model

In order to realize our simulations, we exploited the software NetLogo (Wilensky, 1999).

In every scenario there are four general categories, called Cat1, Cat2, Cat3 and Cat4, composed by 100 agents per category. Each category is characterized by:

1. an **average value of trustworthiness**, in range [0,100];
2. an **uncertainty value**, in range [0,100]; this value represents the interval of trustworthiness in which the agents can be considered as belonging to that category.

These two values are exploited to generate the **objective trustworthiness** of each agent, defined as *the probability that, concerning a specific kind of required information, the agent will communicate the right information*.

Of course the trustworthiness of categories and agents is strongly related to the kind of requested information/task. Nevertheless, for the purpose of our it is enough to use just one kind of information (defined by τ) in the simulations. The categories' trustworthiness of Cat1, Cat2, Cat3 and Cat4 are fixed respectively to 80, 60, 40 and 20% for τ. What changes through scenarios is the uncertainty value of the categories: 1, 20, 50, and 80%.

We want to present a series of scenarios with different settings and referred to localized and non-localized worlds, to show when it is more convenient to exploit recommendations about categories rather than recommendations about individuals, and vice versa.

Both the simulations are composed by two main steps that are repeated continuously. In the first step, called **exploration phase**, agents without any knowledge about the world start experiencing other agents, asking to a subset of the population for the information P. Then they memorize the performance of each queried agent both as individual element and as a member of its own category.

The performance of a agent can assume just the two values 1 or 0, with 1 meaning that the agent is supporting the information P and 0 meaning that it is opposing to P. For sake of simplicity, we assume that P is always true.

The exploration phase has a variable duration, going from 100 ticks to 1 tick. Depending on this value, agents will have a better or worse knowledge of the other agents.

Then, in a second step (**querying phase**) we introduce in the world a trustor (a new agent with no knowlegde about the trustworthiness of other agents and categories, and that has the necessity to trust someone reliable for a given informative task: in our case τ). It will select a given subset of the population and it will query them. In particular, the trustor will ask them for the best category and the best trustee they have experienced.

In this way, the trustor is able to collect information about both the best recommended category and agent.

It is worth underling that the trustor collects information from the agents considering them as equally trustworthy with respect to the task of "providing recommendations". Otherwise it should weigh differently these recommendations. In practice our agents are sincere.

Then it will select an agent belonging to the best recommended category and it will compare it, in terms of objective trustworthiness, with the best recommended individual agent (trustee).

The possible **outcomes** are:

- **trustee wins (t_win)**: the trustee selected with individual recommendation is better than the one selected by the means of category; then this method gets one point;
- **category wins (c_win)**: the trustee selected by the means of category is better than the one selected with individual recommendation; then this method gets one point;
- **equivalent result**: if the difference between the two trustworthiness values is not enough (it is under a threshold), we consider it as indistinguishable result. In particular, we considered the threshold of 3% as, on the basis of previous test simulations, it has resulted a resonable value.

These two phases are repeated 500 times for each setting.
In particular, we will represent this value:

$$\frac{c\_win}{c\_win + t\_win} \qquad (13)$$

This ratio shows how much categories' recommendation is useful if compared to individual recommendation.

Simulations' results are presented in a graphical way, exploiting 3D shapes to represent all the outcomes. These shapes are divided into two area and represented with two different colors:

- the part over 0.5, represented in light gray, in which prevails the category recommendation;
- the one below 0.5, represented in dark gray, in which prevails the individual recommendation.

These graphs represent a useful view about the utility of the categorial role in the different interactional and social contexts.

For each value of uncertainty, we explored 40 different settings, considering all the possible couple of **exploration phase** and **queried trustee percentage**, where:

- exploration phase $\in$ {all-in,1,3,5,10,25,50,100};
- queried trustees' percentage $\in \{5,10,25,50,100\}$.

When the exploration phase assume the value "all-in" the exploration lasts just 1 tick and in that tick every agent experiences all the others. Although this is a limit case, very unlikely in the real world, it is really interesting as each agent has not a good knowledge of the other agent as individual elements (it has experienced them just one time), but it is able to get a really good knowledge of their categories, as it has experienced them as many times as the number of agents for each category. So this is an explicit case in which the recommendations of the agents about categories are surely more informative than the ones about individuals.

First simulation: non-localized world

As previously said, in the first simulation we explore the case in which the communication in the world is not limited by the phisical distance, like in the web context.

Here we will have that:

1. concerning the exploration phase, agents will ask for information P to a random 3% of the population;
2. concerning the querying phase, the trustor will select (again in a random way) a given subset of the population, going from 100% to 5%;
3. in the end, the trustor will select a random member of the most recommended category, to compare it with the most recommended agent.

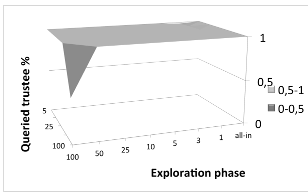

Figure 1.a

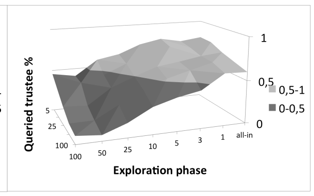

Figure 1.b

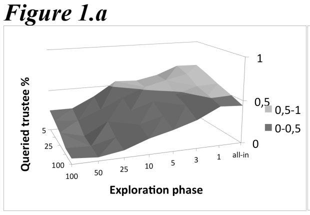

Figure 1.c

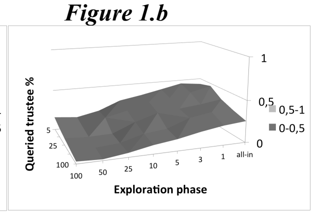

Figure 1.d

Figure 1.a, 1.b, 1.c and 1.d stand respectively for 1%, 20%, 50% and 80% of categories' uncertainty

Second simulation: localized world

Conversely from the previous one, in this simulation everything is ruled by phisical distance:

1. in the exploration phase, on each tick agents move into the world with a probability of 10%; this has the purpose of creating a localization phenomena; then agents will ask for information P to the other trustees which distance in less than 3 NetLogo patches; empirically, we saw that on average they select the 3% of the population, like in the first simulation;
2. in the querying phase, given a percentage of population going from 100% to 5%, the trustor will select the first neighbors until it reaches the requested percentage;
3. in the end, the trustor will select the nearest member of the most recommended category, to compare it with the most recommended agent.

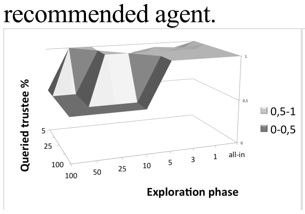

Figure 2.a

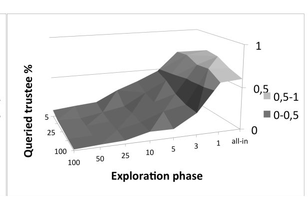

Figure 2.b

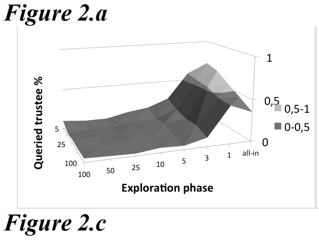

Figure 2.c

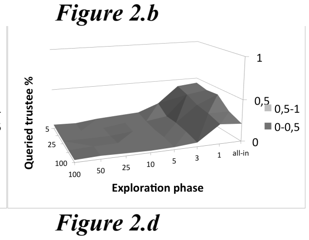

Figure 2.d

Figure 2.a, 2.b, 2.c and 2.d stand respectively for 1%, 20%, 50% and 80% of categories' uncertainty

Results' discussion

Effects in each simulation

Starting the analysis from the common features in the outcome of the two different worlds, we identified three effects. The *first effect* is due to categories' uncertainty: the less it is, the more is the utility of using categories; the more it is, the less categories will be useful. It is possible to notice it looking at the overal picture: the curves of the graphs lower, going from a maximal value in **Figure 1.a and 2.a** to a minimal value in **Figure 1.d and 2.d**. Concretely, one could deal with classes whose members perform accordingly to it, or classes where there is a very high variance: our evaluation on a member of a that category becomes more inaccurate. Because of that, the category's utility decreases.

The *second effect* is due to the exploration phase. The longer this phase is the more individual recommendations are useful; the less it lasts the more category recommendations are useful. This second effect can be described with the fact that each agent, reducing the number of interactions in the explorative phase, will have relevantly less information with respect to the individual agents. At the same time its knowledge with respect to categories does not undergo a significant decline given that categories' performances derive from several different agents.

The *third effect* is introduced by the queried trustee percentage, that acts exactly as the exploration phase: the higher the percentage of queried agents, the more individual's recommendations are useful; the less it is, the more categories' recommendations are useful. In words, reducing the number of queried trustees, the trustor will receive with decreasing probability information about the more trustworthy individuals in the domain, while information on categories maintains a good level of stability, showing a greater robustness.

The exploration phase length and the queried agents percentage cooperate in determining respectively the degree of knowledge (or ignorance) in the world and the level of inquire about this knowledge. In particular, with "the knowledge in the world" we intend how the agents can witness the trustworthiness of the other agents or their aggregate, given the constraints defined by the external circumstances (number and kind of interactions, kind of categories, and so on).

In practice, both these elements seem to suggest that the role of categories becomes relevant when the knowledge within the analyzed system either decreases or degrades (before the interaction with the trustor) or the transferred knowledge (to the trustor) is reduced. In these cases it is better to rely on the categorial recommendations rather than individual ones.

This result reaches the point of highest criticality in the "all-in" case in which, as expected, the relevance of categories reaches its maximal value.

Localized World versus Non-localized World

Let's then discuss the main point of this paper, i.e. the difference between these two main settings: the localized world (L) and non-localized world (NL).

The *first difference* resides in the behavior. While the NL tends to have a convex behavior, the L one tends to be concave: the descent of the categories' utility in the first case is less steep than in the second. The *second effect* is easier to notice: the curves of NL case are quite always higher than the L case.

Both these effects are symptoms of the fact that the utility of categories is higher in the NL case. In fact, in the NL world the agents can have access to more other agents, as they are not constrained by physical distance. In this way, they know more agents, but their knowledge about each single one is limited.

Conversely in the L world, each agent can just query its neighbors. Although they move into the world, their knowledge is strictly related to their physical position. As a consequence, they will know better their neighbors and their knowledge of categories strongly depends on the individuals they have met.

Conclusion

Other works (Falcone et al, 2013; Burnett et al, 2010) show the advantages of using reasoning about categorization to select trustworthy agents. In particular, how it were possible to attribute to a certain unknown agent, a value of trustworthiness with respect to a specific task, on the basis of its classification in, and membership to, one (/or more) category/ies. In practice, the role of generalized knowledge has proven to determine the possibility to anticipate the value of unknown agents.

In this paper we investigated the different roles that recommendations about individual agents and about categories of agents can play, in L and NL worlds.

We showed cases in which categories information is more useful that information towards individual agents, inquiring and matching different dimensions and parameters. Our results show that the information on categories is more robust to knowledge degradation, losing its value more slowly with respect to information about individuals. Moreover we showed that categorial knowledge is considerably more important in NL context, such us the web one, rather than L context.

This analysis can be particularly relevant to decide how to built the cognitive approach of agents searching information among multiple sources. Before choosing between direct or generalized information, we have to evaluate how information is distributed among the agents in the specific domain.

Acknowledgment

This work is partially supported by the project CLARA—CLoud plAtform and smart underground imaging for natural Risk Assessment, funded by the Italian Ministry of Education, University and Research (MIUR-PON).

References

Adomavicius, G., Tuzhilin, A. (2005) Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering (TKDE) 17, 734–749.

Burnett, C., Norman, T., and Sycara, K. 2010. Bootstrapping trust evaluations through stereotypes. In Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'10). 241248.

Burnett, C., Norman, T., and Sycara, K. 2010. (2013) Stereotypical trust and bias in dynamic multiagent systems. ACM Transactions on Intelligent Systems and Technology (TIST), 4(2):26, 2013.

Castelfranchi C., Falcone R., (2010) Trust Theory: A Socio-Cognitive and Computational Model, John Wiley and Sons, April 2010.

Conte R., and Paolucci M., 2002, Reputation in artificial societies. Social beliefs for social order. Boston: Kluwer Academic Publishers.

Falcone R, Castelfranchi C, (2008) Generalizing Trust: Inferencing Trustworthiness from Categories. In Proceedings, pp. 65 - 80. R. Falcone, S. K. Barber, J. Sabater-Mir, M. P. Singh (eds.). Lecture Notes in Artificial Intelligence, vol. 5396. Springer, 2008.

Falcone R., Piunti, M., Venanzi, M., Castelfranchi C., (2013), From Manifesta to Krypta: The Relevance of Categories for Trusting Others, in R. Falcone and M. Singh (Eds.), ACM Transaction on Intelligent Systems and Technology, Volume 4 Issue 2, March 2013

Fang H., Zhang J., Sensoy M., and Thalmann N. M. (2012) A generalized stereotypical trust model. In Proceedings of the 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), pages 698–705, 2012.

Guo G., Zhang J. and Yorke-Smith N., 2014, Leveraging Multiviews of Trust and Similarity to Enhance Clustering-based Recommender Systems, Knowledge-Based Systems, accepted, 2014

Huynh, T.D., Jennings, N. R. and Shadbolt, N.R. , 2006, An integrated trust and reputation model for open multi-agent systems. Journal of Autonomous Agents and Multi-Agent Systems, 13, (2), 119-154.,

Lops P., Gemmis M., and Semeraro G., (2011), "Content-based recommender systems: State of the art and trends," in Recommender Systems Handbook. Springer, pp. 73–105.

Ramchurn S., Jennings N., Sierra C., and Godo L. (2004) Devising a trust model for multi-agent interactions using confidence and reputation. Applied Artificial Intelligence, 18(9-10):833-852.

Sabater-Mir, J. 2003. Trust and reputation for agent societies. Ph.D. thesis, Universitat Autonoma de Barcelona.

Sensoy M., Yilmaz B., and Norman T. J. 2014, STAGE: Stereotypical trust assessment through graph extraction. Computational Intelligence.

Wilensky, U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Yolum, P. and Singh, M. P. 2003. Emergent properties of referral systems. In Proceedings of the 2nd International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS'03).