

Chapter 25

A formal analysis of trust and distrust relationships in Shakespeare's Othello

Elisabetta Erriquez, Wiebe Van der Hoek and Michael Wooldridge

Abstract In settings where agents can be exploited, trust and reputation are key issues. As a consequence, within the multi-agent systems community, considerable research has gone into the development of computationally viable models of these concepts. In this paper, we present an abstract framework that allows agents to form coalitions with agents that they believe to be trustworthy. In contrast to many other models, we take the notion of *distrust* to be our key social concept. We use a graph theoretic model to capture the distrust relations within a society, and use this model to formulate several notions of mutually trusting coalitions. We investigate principled techniques for how the information present in our distrust model can be aggregated to produce individual measures of how trustworthy an agent is considered to be by a society. We extend the framework with *trust* relationships and introduce the notion of *Trust-Coalitions (T-Coalitions)*. We finally provide an analysis of the Shakespeare's tragedy, Othello, using our abstract framework.

1 Introduction

The goal of coalition formation is typically to form robust, cohesive groups that can cooperate to the mutual benefit of all the coalition members. In open distributed systems, where there are many components that can enter and leave the system as they wish, the notion of *trust* becomes key when it comes

Elisabetta Erriquez
Department of Computer Science, University of Liverpool, UK
e-mail: E.Erriquez@liverpool.ac.uk

Wiebe Van der Hoek,
Department of Computer Science, University of Liverpool, UK
e-mail: Wiebe.Van-Der-Hoek@liverpool.ac.uk

Michael Wooldridge
Department of Computer Science, University of Oxford, UK
e-mail: mjw@cs.ox.ac.uk

to decisions about which coalitions to form, and when. When such systems are inhabited by agents that encompass some level of autonomy, each representing their own stakeholders with their own objectives, it not only becomes plausible that some agents are not trustable, the consequences of joining a coalition of which some members cannot be trusted, or do not trust each other, becomes a key aspect in the decision of whether or not to join a group of agents.

With a relatively small number of exceptions [2, 10], existing models of coalition formation do not generally consider trust. In more general models [12, 9], individual agents use information about reputation and trust to rank agents according to their level of trustworthiness. Therefore, if an agent decides to form a coalition, it can select those agents he reckons to be trustworthy. Or, alternatively, if an agent is asked to join a coalition, he can assess his trust in the requesting agent and decide whether or not to run the risk of joining a coalition with him. We argue that these models lack a *global* view. They only consider the trust binding the agent starting the coalition and the agents receiving the request to join the coalition. Here, we address this restriction. We propose an abstract framework through which autonomous, self-interested agents can form coalitions based on information relating to trust. In fact, we use *distrust* as the key social concept in our work. Luckily, in many societies, trust is the norm and distrust the exception, so it seems reasonable to assume that a system is provided with information of agents that distrust each other based on previous experiences, rather than on reports of trust. Moreover, in several circumstances, it makes sense to assume that agents base their decision on which coalition they form on explicit information of distrust, rather than on information about trust. So, we focus on how distrust can be used as a mechanism for modelling and reasoning about the reliability of others, and, more importantly, about how to form coalitions that satisfy some stability criteria. We present several notions of mutually trusting coalitions and define different measures to aggregate the information presented in our distrust model.

Taking distrust as the basic entity in our model allows us to benefit in the sense of deriving our core definitions from an analogy with a popular and highly influential approach within *argumentation theory* [13]. Specifically, the distrust-based models that we introduce are inspired by the *abstract argumentation frameworks* proposed by Dung [4]. In Dung's framework, an attack relation between arguments is the basic notion, which inspired us to model a distrust relation between agents. We show that several notions of stability and of extensions in the theory of Dung naturally carry over to a system where distrust, rather than attack, is at the core. We extend and refine some of these notions to our trust setting. We also introduce an extension of the model, allowing the use of explicitly defined *trust relationships* and we define a concept of coalition based on the trust relations. Finally, we use our model to analyse the (dis)trust relationships in Shakespeare's famous play, *Othello*, giving an actual application and analysis of the models proposed.

Section 2 gives the formal definition of the framework presented and introduces several different notions. Section 2 explains how the information presented in abstract trust frameworks can be *aggregated* to provide a single measure of how trustworthy individuals within the society are. Section 3

presents an extension of the framework introducing also *trust* relationships. Section 4 provides an analysis of the model applied to the Othello's scenario. Finally, Section 5 concludes the paper and presents some possible avenues for future work.

2 Abstract Trust Frameworks

In this section, we introduce the basic models that we use throughout the remainder of the paper.¹ Our approach was inspired by the abstract argumentation frameworks of Dung [4]. Essentially, Dung was interested in trying to provide a framework that would make it possible to make sense of a domain of discourse on which there were potentially conflicting views. He considered the various conflicting views to be represented in *arguments*, with an *attack relation* between arguments defining which arguments were considered to be inconsistent with each other. This attack relation is used to determine which sets of arguments are acceptable in some specified way. Notice that Dung's framework focuses on capturing the notion of two arguments being incompatible, or inconsistent with one-another, rather than supporting each other. In our work, we use similar graph like models, but rather than arguments our model is made up of agents, and the binary relation (which is used in determining which coalitions are acceptable), is a *distrust* relation.

We assume that agents have some incentive for sharing their evaluations of the other agents in the community. Although much previous work deals with trust relationships, in our approach, we consider only *distrust* relationships between the agents. In real life, people do not always share their positive evaluation about others, but they are more inclined to report bad experiences, as a warning to other people and as a way to affect the reputation of the person the bad experience was with, as showed by the large research around *negative word of mouth* [17].

A *distrust* relation between agent i and agent j is intended as agent i having none or little trust in agent j . More precisely, when saying that agent i distrusts agent j we mean that, in the context at hand, agent i has insufficient confidence in agent j to share membership with j in one and the same coalition.

When the agents share their evaluation about others, we can build a distrust network of agents linked by *distrust* relationships.

The follow definitions characterize our formal model.

Definition 1 An Abstract Trust Framework (ATF), S , is a pair:

$$S = \langle Ag, \rightsquigarrow \rangle$$

where:

- Ag is a finite, non-empty set of agents; and
- $\rightsquigarrow \subseteq Ag \times Ag$ is a binary distrust relation on Ag .

¹ An initial version of the framework appeared in [6].

When $i \rightsquigarrow j$ we say that agent i distrusts agent j . We assume \rightsquigarrow to be irreflexive, i.e., no agent i distrusts itself. Whenever i does not distrust j , we write $i \not\rightsquigarrow j$. So, we assume $\forall i \in Ag, i \not\rightsquigarrow i$. Call an agent i fully trustworthy if for all $j \in Ag$, we have $j \not\rightsquigarrow i$. Also, i is trustworthy if for some $j \neq i, j \not\rightsquigarrow i$ holds. Conversely, call i fully trusting if for no $j, i \rightsquigarrow j$. And i is trusting if for some $j \neq i, i \not\rightsquigarrow j$.

Later, we will find it convenient to compare abstract trust frameworks, and for this we use the following definition.

Definition 2 If $S_1 = \langle Ag_1, \rightsquigarrow_1 \rangle$ and $S_2 = \langle Ag_2, \rightsquigarrow_2 \rangle$ are two ATFS, we say that S_2 extends S_1 , written $S_1 \sqsubseteq S_2$, if both $Ag_1 \subseteq Ag_2$ and $\rightsquigarrow_1 \subseteq \rightsquigarrow_2$.

Coalitions with Trust

In what follows, when we refer to a “coalition” it should be understood that we mean nothing other than a subset C of Ag . When forming a coalition, there are several ways to measure how much distrust there is among them, or how trustable the coalition is with respect to the overall set of agent Ag .

Definition 3 Given an ATF $S = \langle Ag, \rightsquigarrow \rangle$, a coalition $C \subseteq Ag$ is distrust-free if no member of C distrusts any other member of C .

Note that the empty coalition \emptyset and all singleton coalitions $\{i\}$ are distrust-free: we call these trivial coalitions.



Fig. 1 S_1 , an example of ATFS

Distrust freeness can be thought of as the most basic requirement for a *trusted* coalition of agents. It means that a set of agents has no internal distrust relationships between them. Since we assume \rightsquigarrow to be irreflexive, we know that for any $i \in Ag$, the coalition $\{i\}$ is distrust-free, as is the empty coalition. A distrust-free coalition for S_1 in Figure 1 is, for example, $\{a, c, d\}$.

Consider ATF S_1 from Figure 1. The coalition $C_1 = \{c, d\}$ is distrust-free, but still, they are not angelic: one of their members is being distrusted by some agent in Ag , and they do not have any justification to ignore that. Compare this to the coalition $C_2 = \{a, c, d\}$: any accusations about the trustworthiness of c by b can be neutralised by the fact that a does not trust b in the first place. So, as a collective, they have a defense against possible distrust against them.

With this in mind, we define the following concepts.

Definition 4 Let ATF $S = \langle Ag, \rightsquigarrow \rangle$ be given.

- An agent $i \in Ag$ is called *trustable with respect to a coalition* $C \subseteq Ag$ iff $\forall y \in Ag((y \rightsquigarrow i) \Rightarrow \exists x \in C(x \rightsquigarrow y))$.

- A coalition $C \subseteq Ag$ is a trusted extension of S iff C is distrust-free and every agent $i \in C$ is trustable with respect to C .
- A coalition $C \subseteq Ag$ is a maximal trusted extension of S if C is a trusted extension, and no superset of C is one.

It is easy to see that if $i \in Ag$ is trustable with respect to some coalition C , then i is also trustable with respect to any bigger coalition $C' \supseteq C$. We will see that (maximal) trusted extensions are not closed under supersets, though.

Personal Extensions

In large societies, it is very unlikely that a single agent manages to interact with everyone in the society. For this reason, it has to rely on information given by others, about reputation of the agents it doesn't know. Reputation can be defined as the opinion or view of someone about something [?]. This view can be mainly derived from an aggregation of opinions of members of the community about one of them. However, it is possible that the agent does not trust a particular agent and it wants to discard its opinion. Therefore, when it comes to forming a coalition, the agents wants to consider only its personal opinion and the opinion of the agent it trusts, while still keeping the coalition distrust-free.

For example, suppose that an agent wants to start a project and it needs to form a coalition to achieve its goals. It wants to form a coalition composed only of agents it trusts and who have no distrust relations among them. To capture this intuition, we introduce the notions of *unique personal extension*, which make it precise ².

Definition 5 Given an ATF $S = \langle Ag, \rightsquigarrow \rangle$, and an agent $a \in Ag$, the unique personal extension $UPE(S, a)$, we require, has the following properties:

1. $a \in UPE(S, a)$
2. $UPE(S, a)$ is unique
3. $UPE(S, a)$ is distrust free
4. there is a minimal set $OUT \subseteq Ag$, with the following properties, for all $x, y \in Ag$:
 - a. $x \rightsquigarrow a \Rightarrow x \in OUT$
 - b. $(y \in UPE(S, a) \ \& \ y \rightsquigarrow x) \Rightarrow x \in OUT$
 - c. $y \in UPE(S, a) \Leftrightarrow \forall z(z \rightsquigarrow y \Rightarrow z \in OUT)$

Loosely put: we add a to $UPE(S, a)$, and then we ensure that whoever distrusts or is distrusted by somebody in $UPE(S, a)$ is out, while $UPE(S, a)$ only accepts those agents as members that are at most distrusted by members of OUT . For a more detailed algorithm, see [6].

Note that agents can be out for two reasons: first of all, they may distrust agent a , or they may themselves be distrusted by an agent that is in.

Consider the ATF $S1 = \langle Ag, \rightsquigarrow \rangle$ in Figure 1. Suppose agent a wants to compute its personal extensions. Then, according to our definition, the *unique trusted*

² For more details please refer to [6]

extension computed is $\{a, c, d\}$. Now, suppose agent b wants to compute its personal extensions. Then, according to our definition, agents a will be discarded because it distrust b . Therefore, the *unique trusted extension* for agent b is $\{b, d\}$.

Aggregate Trust Measures

Abstract trust frameworks provide a social model of (dis)trust; they capture, at a relatively high level of abstraction, who (dis)trusts who in a society, and notions such as trusted extensions and personal extensions use these models to attempt to understand which coalitions are free of negative social views. An obvious question, however, is how the information presented in abstract trust frameworks can be *aggregated* to provide a single measure of how trustworthy (or otherwise) an individual within the society is. We now explore this issue. We present three aggregate measures of trust, which are given relative to an abstract trust framework $S = \langle Ag, \rightsquigarrow \rangle$ and an agent $i \in Ag$. Both of these trust values attempt to provide a principled way of measuring the overall trustworthiness of agent i , taking into account the information presented in S :

- *Distrust Degree:*
This value ignores the structure of an ATF, and simply looks at how many or how few agents in the society (dis)trust an agent.
- *Expected trustworthiness:*
This value is the ratio of the number of maximal trusted extensions of which i is a member to the overall number of maximal trusted extensions in the system S .
- *Coalition expected trustworthiness:*
This value attempts to measure the probability that an agent $i \in Ag$ would be trusted by an arbitrary coalition, picked from the overall set of possible coalitions in the system.

These latter two values are related to solution concepts such as the Banzhaf index, developed in the theory of cooperative games and voting power, and indeed they are inspired by these measures [8].

Distrust Degree

On the web, several successful approaches to credibility such as PageRank [3, 11] use methods derived from graph theory to model credibility, which utilize the connections of the resource for evaluation. Several graph theoretic models of credibility and text retrieval [15] rely on the consideration of the in-degree of the vertex, that is the sum of the incoming edges of that particular vertex in a directed graph. The degree of the incoming edges is used to extract importance and trustworthiness.

In our model, incoming edges are distrust relationships, therefore they represent a negative evaluation of a particular agent from the others in the society. Thus, measuring the in-degree of an agent in the society can give an indication how reliable (or unreliable) that agent is considered overall.

Formally, we call this value the *distrust-degree* for an abstract trust framework $S = \langle Ag, \rightsquigarrow \rangle$ and an agent $i \in Ag$, denoted as $\delta_i(S)$, and it is defined:

$$\delta_i(S) = \frac{|\{x \mid x \in Ag \text{ and } x \rightsquigarrow i\}|}{|Ag|}.$$

This number provides us a measure of the reliability of the agent in the whole society. The higher the number of agents that distrust it, the less reliable that agent is considered to be.

However, as we mentioned before, a maximal trusted extension or, in general, a coalition C , according to our approach, is a set of agents who trust each other. Therefore, these agents may not be interested in the evaluation of the agents outside the coalition. They are more interested in a distrust degree relative to C . Hence, we define the following measure. The *coalition distrust-degree* for an abstract trust framework $S = \langle Ag, \rightsquigarrow \rangle$, a coalition C and an agent $i \in Ag$, denoted as $\delta_i^C(S)$, defined as:

$$\delta_i^C(S) = \frac{|\{x \mid x \in C \text{ and } x \rightsquigarrow i\}|}{|C|}.$$

The coalition distrust degree provides a measure for the agents in C to select agents outside the trusted coalition, who they believe to be more reliable among the agents in the society. Agents in C can rank the agents outside using the value of the coalition distrust degree. In this way, it is possible to obtain an ordered list of the agents who the coalition consider less unreliable. The smaller the value of the coalition distrust-degree, the more reliable the agent is considered.

Expected Trustworthiness

As we noted above, the expected trustworthiness of an agent i in system S is the ratio of the number of maximal trusted extensions in S of which i is a member to the overall number of maximal trusted extensions in the system S . To put it another way, this value is the probability that agent i would appear in a maximal trusted extension, if we picked such an extension uniformly at random from the set of all maximal trusted extensions. Formally, letting $mte(S)$ denote the set of maximal trusted extensions in $S = \langle Ag, \rightsquigarrow \rangle$, the expected trustworthiness of agent $i \in Ag$ is denoted $\mu_i(S)$, defined as:

$$\mu_i(S) = \frac{|\{C \in mte(S) \mid i \in C\}|}{|mte(S)|}.$$

Clearly, if $\mu_i(S) = 1$ then i is strongly trusted, according to the terminology introduced above, and moreover a is weakly trusted iff $\mu_i(S) > 0$.

From existing results in the argumentation literature on computing extensions of abstract argument systems [5], we can also obtain the following:

As an aside, note that the expected trustworthiness value is inspired by the *Banzhaf index* from cooperative game theory and voting theory [8].

Coalition Expected Trustworthiness

There is one obvious problem with the overall expected trustworthiness value, as we have introduced above. Suppose we have a society that is entirely trusting (i.e., the entire society is distrust free) apart from a single “rogue” agent, who distrusts everybody apart from himself, even though everybody trusts him. Then, according to our current definitions, there is no maximal trusted extension apart from the rogue agent. This is perhaps counter intuitive. To understand what the problem is, observe that when deriving the value $\mu_i(S)$, we are taking into account the views of *all* the agents in the society – which includes every rogue agent. It is this difficulty that we attempt to overcome in the following measure. To define this value, we need a little more notation. Where $R \subseteq X \times X$ is a binary relation on some set X and $C \subseteq X$, then we denote by $\text{restr}(R, C)$ the relation obtained from R by restricting it to C :

$$\text{restr}(R, C) = \{(s, s') \in R \mid \{s, s'\} \subseteq C\}.$$

Then, where $S = \langle Ag, \rightsquigarrow \rangle$ is an abstract trust framework, and $C \subseteq Ag$, we denote by $S \downarrow C$ the abstract trust framework obtained by restricting the distrust relation \rightsquigarrow to C :

$$S \downarrow C = \langle C, \text{restr}(\rightsquigarrow, C) \rangle.$$

Given this, we can define the *coalition expected trustworthiness*, $\varepsilon_i(S)$, of an agent i in given an abstract trust framework $S = \langle Ag, \rightsquigarrow \rangle$ to be:

$$\varepsilon_i(S) = \frac{1}{2^{|Ag|-1}} \sum_{C \subseteq Ag \setminus \{i\}} \mu_i(S \downarrow C \cup \{i\}).$$

Thus, $\varepsilon_i(S)$ measures the expected value of μ_i for a coalition $C \cup \{i\}$ where $C \subseteq Ag \setminus \{i\}$ is picked uniformly at random from the set of all such possible coalitions. There are $2^{|Ag|-1}$ coalitions not containing i , hence the first term in the definition.

The coalition expected trustworthiness value arguably gives a clearer overall idea of what the trustworthiness of an agent would be with respect to the maximal trusted extensions that can potentially be formed, therefore it offers a better insight in the trust issue related to the problem of forming coalitions.

3 A Framework for trust and distrust . . . An extension

In this section, we present an extension of the AFT described in Section 2. As mentioned in Section 2, in the AFT, the *absence of a distrust relationship* between two agents does not necessarily imply the *presence of a trust relationship*. Therefore it seems natural to complement the distrust framework just introduced with trust relationships. The follow definitions characterize our formal model.

Definition 6 An Abstract Trust/Distrust Framework (ATDF), S , is a triple:

$$S = \langle Ag, \rightsquigarrow, \dashrightarrow \rangle$$

where:

- Ag is a finite, non-empty set of agents; and
- $\rightsquigarrow \subseteq Ag \times Ag$ is a binary distrust relation on Ag , as described in Section 2.
- $\dashrightarrow \subseteq Ag \times Ag$ is a binary trust relation on Ag .
- $\forall i, j \in Ag$ if $i \rightsquigarrow j$ then $i \not\rightarrow j$;
- $\forall i \in Ag : i \dashrightarrow i$;

When $i \dashrightarrow j$ we say that agent i trusts agent j . We assume that there always exist an implicit trust relationship from and to the agent itself. Whenever i does not trust j , we write $i \not\rightarrow j$.

In this case, as well as before, the absence of an explicit *trust* relationship does not imply the presence of a *distrust* relationship. Note that the fourth bullet of the definition state that an agent cannot trust and distrust another agent at the same time. However we do not impose to any of the relationships to be mutual, therefore there is no restriction to prevent a situation where for two agents i, j , $i \rightsquigarrow j$ and $j \dashrightarrow i$, as shown in the situation between agent a and b in S_2 in Figure 2.

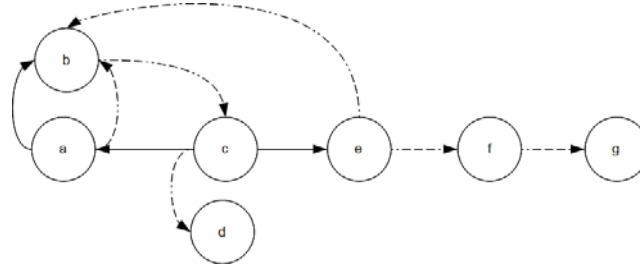


Fig. 2 S_2 , an example of ATDFS

We now present the concept of coalition based on *trust relationship*.

Definition 7 Given an ATDF $S = \langle Ag, \rightsquigarrow, \dashrightarrow \rangle$, a coalition $C \subseteq Ag$ is a T-Coalition (Trust-Coalition) if:

1. C is distrust-free wrt the \rightsquigarrow relationship;
2. $\forall i \in C, \exists j \in C : i \dashrightarrow j$ or $j \dashrightarrow i$;
3. C is maximal wrt \subseteq among the sets satisfying 1 and 2.

Note that the singleton coalitions $\{i\}$ are *T-Coalitions* as there is always an implicit *trust* relationship from and to the agent itself: we call them trivial T-Coalitions.

In S_2 , shown in Figure 2, the *T-Coalitions*, apart from the trivial ones, are $\{b, c, d\}$ and $\{b, e, f, g\}$.

4 Othello, An Analysis

In this section we use the famous tragedy by Shakespeare, *Othello*, to give an actual illustration and analysis of the models proposed. The play revolves around four central characters: Othello, a Moorish general in the Venetian army; his wife Desdemona; his lieutenant, Cassio; and his trusted ensign Iago. Othello is a general in the Venetian Republic military, who is in command of the army fighting the Turks attacking Cyprus.

The play opens with Roderigo, a rich and dissolute gentleman, complaining to Iago, a highly-ranked soldier, that Iago has not told him about the secret marriage between Desdemona, the daughter of a Senator named Brabantio, and Othello. He is upset by this development because he loves Desdemona and had previously asked her father for her hand in marriage. Iago is upset with Othello for promoting Cassio, a younger man, above him, and tells Roderigo that he plans to use Othello for his own advantage. Iago tries in many ways to have Cassio stripped off his rank but he also wants to ruin Othello for having preferred Cassio over him. Therefore Iago persuades Othello to be suspicious of Cassio and Desdemona. He achieves this using a precious handkerchief that was Othello's first gift to Desdemona and giving it to Cassio, with the help of his wife Emilia. The fake defense from Iago towards Cassio and his deliberate reticence are the central part of Iago's work of persuasion which leads to Othello killing Desdemona in a fit of rage. In the epilogue, Emilia reveals that Desdemona's betrayal was invented by Iago, who immediately kills her. Othello, feeling guilty for killing the innocent Desdemona, kills himself. Iago is arrested and Cassio takes Othello's place as general.

The play is divided in five acts. We take this division as the natural breaking points for modelling the scenarios with our AFT and AFDT . We design an AFT and AFDT for each act, using the main characters and the \rightsquigarrow and \dashrightarrow relationships, representing the distrust (or more in general dislike) and trust relationships among the characters.

For each of the main characters we give the aggregate trust measures for each act and for Othello and Iago, the primary actors, we also present the UPE .

In our AFTS representing the Othello's scenario, we only consider the main characters: Othello, Desdemona, Iago, Cassio, Roderigo and, for the first act, Brabantio. Therefore our context is limited to the relationships among these figures.

Normally, in multi-agent systems scenarios, a coalition is a set of agents who may or may not work together to achieve a common goal or to earn higher utility [18]. In our example, the Othello's scenario, we abstract from the goal of the coalition and we consider the proposed trusted set of agents simply as sets whose components trust, or have no reason to distrust each other. In the next paragraphs we provide a more extensive description of each act [1] and then we provide a detailed analysis of the society.

ACT 1 Synopsis

Shakespeare's famous play of love turned bad by unfounded jealousy begins in Venice with Iago, a soldier under Othello's command arguing with Roderigo, a wealthy Venetian. Roderigo has paid Iago a considerable sum of money to spy on Othello for him, since he wishes to take Othello's girlfriend, Desdemona as his own.

Roderigo fears that Iago has not been telling him enough about Desdemona and that this proves Iago's real loyalty is to Othello rather than him. Iago explains his hatred of Othello for choosing Cassio as his officer or lieutenant and not him as he expected. To regain Roderigo's trust, Iago and Roderigo inform Brabantio, Desdemona's father, of her relationship with Othello the "Moor", which enrages Brabantio into sending parties out at night to apprehend Othello for what in Brabantio's eyes must be an abuse of his daughter by Othello. Iago lies that Roderigo and not himself, was responsible for angering Brabantio against Othello, Iago telling Othello that he should watch out for Brabantio's men who are looking for him. Othello decides not to hide, since he believes his good name will stand him in good stead.

We learn that Othello has married Desdemona. Brabantio and Roderigo arrive, Brabantio accusing Othello of using magic on his daughter. Othello stops a fight before it can happen but Othello is called away to discuss a crisis in Cyprus, much to the anger of Brabantio who wants justice for what he believes Othello has done to his fair Desdemona. The Duke is in council with several senators discussing their enemy, the Turks (Turkish people). Brabantio complains to the Duke that Othello bewitched his daughter and had intimate relations with her. Desdemona is brought in to settle the matter, Othello meanwhile explains how he and Desdemona fell in love. Desdemona confirms this and the Duke advises Brabantio that he would be better off accepting the marriage than complaining and changing nothing.

The Duke orders Othello to Cyprus to fight the Turks, with Desdemona to follow, accompanied by the trusted Iago. Roderigo despairs that his quest for Desdemona is over now that she is married, but Iago tells him not to give up and earn money instead; soon Desdemona will bore of Othello.

Alone, Iago reveals his intention to continue using Roderigo for money and his hatred of Othello (Othello picked Cassio and not Iago for his lieutenant). Iago explains that his plan is avenge Othello is to suggest to Othello that Cassio is sleeping with Desdemona (Othello's wife).

ACT 1 Analysis

In figure 3 and 4 we can observe that the only distrust-free characters, at the end of the first act, are Desdemona and Roderigo.

The *maximal trusted extension* is formed by

$$\{\text{Desdemona}, \text{Roderigo}, \text{Cassio}, \text{Brabantio}\}.$$

Note that Othello, although a noble person at the beginning of our story, is hated and therefore distrusted by Iago, who feels betrayed by him, by Roderigo who feels Desdemona has been stolen from him by Othello, and by Brabantio, Desdemona’s father, who does not approve of the couple’s marriage.

Table 1 Table showing the Maximal Trusted Extensions and the T-Coalitions in the Othello’s society S for each Act

	Maximal Trusted Extensions	T-Coalitions
Act 1	$\{Desdemona, Roderigo, Cassio, Brabantio\}$	$\{Desdemona, Brabantio\}$ $\{Othello, Desdemona, Cassio\}$
Act 2	$\{Desdemona, Roderigo, Iago\}$	$\{Roderigo, Iago\}$ $\{Othello, Desdemona, Cassio\}$
Act 3	$\{Desdemona, Iago\}$	$\{Desdemona, Cassio\}$
Act 4	$\{Desdemona, Iago\}$	$\{Desdemona, Cassio\}$
Act 5	$\{Desdemona, Iago\}$ $\{Desdemona, Cassio\}$ $\{Desdemona, Roderigo\}$	$\{Othello, Desdemona\}$ $\{Cassio, Desdemona\}$

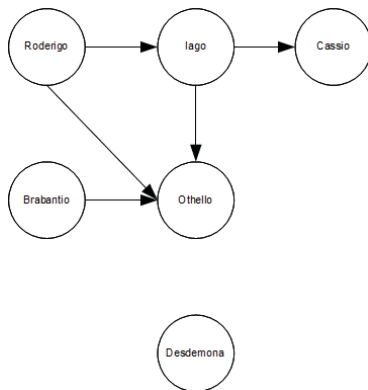


Fig. 3 ATF for Act 1

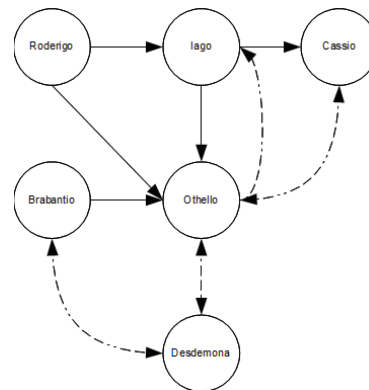


Fig. 4 ATDF for Act 1

Therefore Othello is highly distrusted in the first act, as shown by his *coalition expected trustworthiness* value, 0.125, which is the lowest among all characters in this act. However, his beloved Desdemona and his loyal lieutenant Cassio trust him and, in fact they are part of Othello’s *unique personal extension*. Desdemona,

ACT 2 Synopsis

Several weeks later in Cyprus, Othello's arrival is expected. But a terrible storm has largely battered and destroyed the Turkish fleet, which no longer poses a threat to Cyprus. Unfortunately there are fears that this same storm drowned Othello as well. Many people praise Othello. Cassio, who has arrived, sings Desdemona's praises. A ship is spotted but it is Desdemona and Iago's, not Othello's. Iago suspects that Cassio loves Desdemona and slyly uses it to his advantage. Iago tells Roderigo that he still has a chance with Desdemona but Cassio whom Desdemona could love is in the way. Killing Cassio (who became Othello's lieutenant instead of Iago) will leave Desdemona to Roderigo, Iago slyly explains.

Othello finally arrives, to everyone's great relief. Iago decides to tell Othello that Cassio is having an affair with Desdemona, so Iago will be rewarded whilst Cassio will be punished. A Herald announces celebration that "our noble general Othello!" has defeated the Turkish fleet, calling on all to celebrate this great triumph and also to celebrate Othello's "nuptial" or wedding to the fair Desdemona.

Iago learns more of Cassio's high regard for Desdemona and Iago manipulates Cassio into drinking too much since he is certain Cassio will do something he will regret. Iago also tells Roderigo to attack Cassio. This happens, and Cassio wounds Roderigo. Othello is now awake and Cassio's name ruined. Othello, although he loves Cassio, has no choice but to demote him from his position as his lieutenant. Next Iago comforts Cassio by suggesting he speak with Desdemona who could put in a good word for him with Othello. Iago comforts a wounded Roderigo, telling him he has won by ruining Cassio's name. Iago has his wife Emilia ensure Desdemona and Cassio will talk so Othello can see his wife talking with Cassio, allowing Iago to convince Othello that Desdemona is being unfaithful.

ACT 2 Analysis

In the second act, as shown in Figure 5 and 6, the distrust and trust relationships between the characters change. Our models reflect these changes.

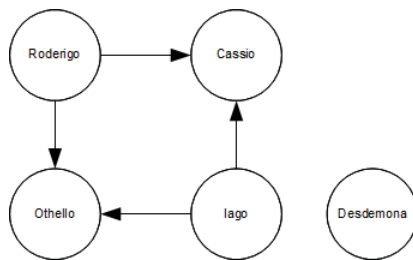


Fig. 5 ATFD for Act 2

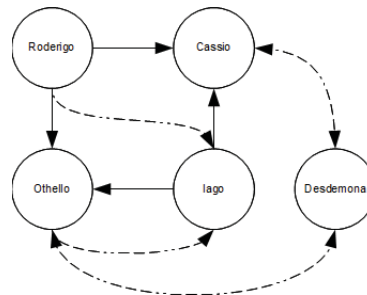


Fig. 6 ATDF for Act 2

We can observe that now Iago has become distrust free, since Roderigo has regained trust in him. Roderigo now distrusts Cassio, as well as Othello, convinced by Iago that he should attack Cassio to ruin Othello's name.

The *Maximal trusted extension* now includes Iago and excludes Cassio. It is, in fact formed by

{Desdemona, Roderigo, Iago}.

This newly acquired distrust-freeness, boosts Iago *coalition expected trustworthiness* to 1, while Cassio's decreases to 0.25, since Othello has lost trust in him because of his fight with Roderigo.

Iago's *unique personal extension* now includes Roderigo, having the conflict being removed.

Othello's *coalition expected trustworthiness* improves slightly, being 0.25 in this act. This is due to the absence of Brabantio in this act. His *unique personal extension* does not vary. It still includes Cassio because, even though Othello has lost *trust* in him, due to his fight, Iago has still not convinced Othello of Cassio's betrayal.

The *T-Coalitions* now include the one formed by {Roderigo, Iago} which reflect the actual cooperation between the two in this second act, even if Iago's motives are not genuine.

ACT 3 Synopsis

Cassio tells Iago that he has arranged to meet Desdemona, Iago helping Cassio to do this. Iago's wife, Emilia, tells Cassio that Othello would like to reinstate him as his lieutenant but the fact that Cassio's fight is public news, prevents Othello from doing this immediately. Emilia tells Cassio that she can arrange a meeting with Desdemona. Some time later, Cassio speaks with a very sympathetic Desdemona who assures him that Othello still very much loves Cassio. Furthermore, Desdemona resolves to keep putting in a good word for Cassio until he is again Othello's lieutenant. At a distance, Iago manipulates Othello by first suggesting shock and then hiding his outbursts from Othello. This guarantees Othello's attention, as Iago plants seeds of doubt in Othello's mind about Desdemona's fidelity especially where Cassio is concerned. Iago leaves Othello almost convinced that his wife is having an affair with Cassio.

Othello now complains of a headache to Desdemona, which results in her dropping a strawberry patterned handkerchief, Othello's first gift to her. Emilia picks this up gives it to Iago who decides the handkerchief could help his manipulation if he ensures Cassio receives it. Iago arranges to place the handkerchief near Cassio's lodgings or home where he is certain to find it and take it as his own, unaware that it is Othello's gift to Desdemona.

A furious Othello returns to Iago, certain his wife is faithful and demanding proof from Iago of Desdemona's infidelity. Reluctantly and hesitantly, Iago tells Othello he saw Cassio wipe his brow with Desdemona's handkerchief. Othello is convinced, cursing his wife and telling Iago who is now promoted to lieutenant to kill Cassio. Othello will deal with Desdemona. Desdemona worries about her missing handkerchief and comments that if she lost it, it could

lead Othello doubting her fidelity. Emilia when asked about Desdemona's lost handkerchief, lies, denying having seen the handkerchief she picked up and gave to Iago. Othello enters; asking Desdemona for the very same handkerchief and Desdemona assures him that the handkerchief is not lost and will be found. Desdemona now tries to change the subject to Cassio, but Othello continually stresses the value the handkerchief has to him, this leading to Othello angrily ordering his wife away.

Cassio arrives, Desdemona telling him that her attempts to help him are not going well. Iago claims total ignorance to the cause of Othello's fury. Cassio gives Othello's handkerchief, which he found, to his suspicious mistress Bianca.

ACT 4 Synopsis

Iago fans the flames of Othello's distrust and fury with Desdemona's supposed "infidelity" by first suggesting Desdemona shared her bed with Cassio and then that her giving away the handkerchief is no big deal when Iago knows exactly how hurtful to Othello, giving away this sentimental gift is.

Next Iago suggests to Othello that Cassio will "blab" or gloat to others about his conquest of Desdemona before telling Othello that Cassio boasted to him that he did indeed sleep with Desdemona. Meeting later with Cassio, Iago cunningly talks to Cassio about Cassio's mistress Bianca, each smile and each gesture made by Cassio infuriating a hidden Othello who thinks Cassio is talking about sleeping with Desdemona (Othello's wife). Next Bianca (Cassio's mistress) arrives, angrily giving back the handkerchief Cassio gave to her. This infuriates Othello since as Iago puts it, Cassio not only received Othello's handkerchief from his wife but then gave it away to his whore (Bianca) as if it were worthless. Othello decides to kill Desdemona by strangulation in her bed, Iago's idea. Iago pledges to kill Cassio.

Lodovico arrives, announcing that Othello is to return home and Cassio is to be the next Governor of Cyprus. Desdemona's joy for Cassio enrages Othello, leaving Lodovico and Iago to wonder how much Othello seems to have changed and leaving poor Desdemona to wonder how she offended the man she truly loves. Othello questions Emilia as to whether Desdemona was unfaithful to him. Annoyed that Emilia's answers suggest nothing has happened between Desdemona and Cassio, Othello dismisses her comments as those of a simple woman. Othello meets Desdemona, Desdemona becoming increasingly upset with her husband's anger towards her, an anger she cannot understand. Othello eventually reveals to Desdemona that her infidelity is the source of his anger, Desdemona pleading her innocence on deaf ears. Emilia and Desdemona discuss Othello's strange behavior. Emilia is certain some evil fellow has twisted Othello to believe Desdemona has been unfaithful, not realizing that this evil man is her own husband Iago.

We learn that Iago has been pocketing Roderigo's gifts to Desdemona, which never reached her. Fearing Roderigo will learn this, Iago tells Roderigo that Cassio must die since Iago benefits if ever man dies. Lodovico tries to calm

Othello down. Othello orders Desdemona to bed to await him later, an order Desdemona dutifully obeys out of love for Othello. Emilia notices that Othello is much calmer now and tells Desdemona her bed has been made with her wedding sheets as requested. Desdemona asks to be buried in those same sheets should she die before Emilia, a hint of trouble ahead.

Emilia is barred from joining Desdemona in her bedchamber, angering her. Desdemona, depressed, recalls a song (The Willow Song) of a maid who was similarly abused by her husband and sings it. Desdemona and Emilia talk about infidelity. Desdemona would not be unfaithful to her husband (Othello) for all the world; the more cynical and worldly Emilia would for the right price.

ACT 3 and 4 Analysis

In the third act, Othello is convinced that Desdemona and Cassio are having an affair, therefore he distrusts them both. Hence, in our model two new *distrust* relationships appear.

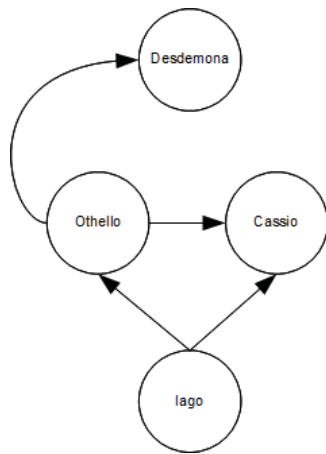


Fig. 7 ATDF for Act 3 and 4

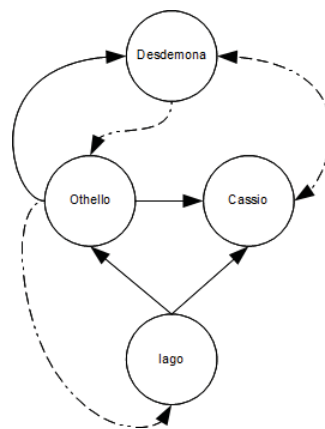


Fig. 8 ATDF for Act 3 and 4

Desdemona, although she is now distrusted by Othello, remains in the *maximal trusted extension*, indirectly defended by the Iago's distrust toward Othello. Iago still remains distrust-free in the society.

Cassio's *coalition expected trustworthiness* decrease to 0.25, while Iago's reaches 1. This reflect the actual situation in the play. Iago is considered a very loyal friend.

Othello is now alone. He does not trust Cassio or Desdemona anymore. This is shown by the fact that his *unique personal extension* is empty, or more accurately that he is the only one in it.

Note that now the only *T-Coalitions* is formed by {Desdemona, Cassio} who, being still unaware of Iago's plot, keep trusting each other.

In the fourth act, the *trust* and *distrust* relationships remain unchanged.

ACT 5 Synopsis

Iago and Roderigo wait in a street to ambush Cassio. Iago tells Roderigo how to kill him. Iago does not care which ends up dead. Iago is worried that about Roderigo's increasing questioning of what happened to jewels that were given to him to pass on to Desdemona. Roderigo attacks Cassio but Cassio wounds Roderigo instead. Iago from behind stabs Cassio, wounding him in the leg. Othello hearing Cassio's cries is pleased, announcing that he too will soon kill (Desdemona). Lodovico and Gratiano (Desdemona's cousin and uncle) and Iago reappear, Iago claiming total innocence to Cassio's injuries even though he inflicted them. Seizing Roderigo, Iago stabs and wounds him "in revenge" for wounding his "friend" Cassio. Gratiano and Lodovico tend to Cassio's wound. Bianca, Cassio's mistress arrives, Iago cleverly laying suspicion for Cassio's injuries on his innocent mistress, making Iago less suspicious.

Othello enters Desdemona's bedchamber (bedroom) trying to convince himself that he is killing her for her own good. He kisses his still asleep wife one last time. Desdemona awakens, but Othello will still kill her, telling her to pray so her soul will not die when she does. Desdemona again asks what wrong she has committed, Othello telling her that she gave Cassio his handkerchief, by which he means he thinks she had an affair with him. Desdemona pleads her innocence, telling Othello to bring Cassio over to prove she did not give away her handkerchief. Othello says he confessed and is dead, Desdemona's fear and surprise prompting Othello to believe she does care for him.

Othello kills Desdemona.

Emilia, banging on the door outside, cannot stop this. Later Emilia is let in, revealing Iago has killed Roderigo and Desdemona who was thought dead, murmurs her last breaths but loyally does not say Othello killed her. Othello tells Emilia he killed her and Emilia despite Iago's attempts to remove her reveals the truth about the handkerchief; she found it, and then gave it to Iago. Iago now in trouble, stabs his wife Emilia and escapes. Emilia dies, singing the "Willow Song" before criticizing Othello for killing his loving wife.

Lodovico, Cassio and the now captured prisoner Iago soon appear, Othello stabbing Iago but not killing him before having his sword removed. Lodovico is disappointed that Othello, a man so honorable has reverted to acting like a slave. Othello tries to argue that killing his wife was a noble action but it falls on deaf ears. Lodovico learns that Othello and Iago plotted Cassio's death. Lodovico reveals letters in the dead Roderigo's pocket proving Cassio was to be killed by Roderigo. Iago proudly confirms that Cassio did find the handkerchief in his bedchamber because Iago placed it there to be found. Othello, realizing what he has done, kills himself with a concealed weapon and lies himself on top of his wife. Cassio is placed in charge of Iago and Lodovico leaves to discuss this sad matter with others abroad.

ACT 5 Analysis

In the end of the final act, the truth is finally revealed. Roderigo finally understands that Iago has been using him. Othello realises that he has been a fool and victim of Iago's plot. This is shown by the new *distrust* relationships from Roderigo and Othello to Iago.

Desdemona is now dead and Cassio realises that Othello was planning to kill him too. Cassio now distrusts Othello and Iago. On the other side, Othello now trusts Cassio again, realizing that Cassio is always been loyal to him.

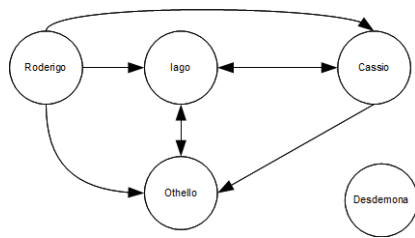


Fig. 9 ATF for Act 5

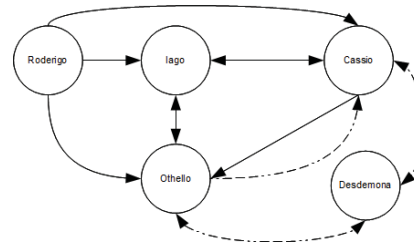


Fig. 10 ATDF for Act 5

The last act changes Iago's situation drastically. His *coalition expected trustworthiness* drops from 1 to 0.3. His *unique personal extension* includes only Desdemona who died before realising she was part of the Iago's plot.

Othello's situation remains bad. His *coalition expected trustworthiness* is very low, 1.18. His *unique personal extension* includes Desdemona, who, even before dying, refused to betray his beloved husband by revealing that he was his killer.

Othello's renewed trust in Desdemona means that there is now another *T-Coalition* formed by the two, since Desdemona never stopped trusting Othello.

5 Conclusions and Future Work

In this paper we have addressed some of the limitations of existing trust-based coalition formation approaches. We have taken the notion of *distrust* to be our key social concept. The main contribution of this work is the definition of an abstract framework that overcomes these limitations allowing the agents to form distrust-free coalitions. We have formulated several notions of mutually trusting coalitions. We have also presented techniques for how the information presented in our distrust model can be aggregated to produce individual measures to evaluate the trustworthiness of the agent with respect to the whole society or to a particular coalition. We also have presented a way to combine the trust and distrust relationships to form coalitions which are still distrust-free.

The analysis of the Othello's scenario shown that the models can successfully reproduce combinations of trust and distrust relationships for realistic situations. The changes in the relationships are reflected in our models and so are the values of the aggregate measures, working as a mirror of the society.

Our model is not utility based, so we are not considering stability from a utility-theoretic point of view, but there is nevertheless an interesting relationship between our notion of stability and that of cooperative game theory. In our view, a coalition is stable if no agent has any rational incentive to distrust any of the members. Future work might consider examining the role of our distrust models in coalition formation in more detail, perhaps in the context of coalition formation algorithms such as those that have recently been proposed within the multi-agent systems community (see, e.g., [14, 16]).

There are many potential directions for future work. We assume that the agents are willing to share their information about the trust they have in other agents. It would be interesting to develop some form of incentives for the agents to do so. Also, a natural development would be to devise a way to integrate the information inferred from the trust relationships into the aggregate trust measures to give a more comprehensive view of the agents status.

Many studies have gone toward trust properties, such as transitivity. As shown in [7], transitivity should be considered in the context of a particular task. In this work, we have abstracted from the particular goal the agents wish to achieve. An interesting development would be to incorporate this information in our concept of mutually trusting coalitions.

Finally, our work is based on a boolean notion of trust. A further improvement would be to add degrees of trust or distrust to allow for more detailed information.

References

1. Othello summary (2005). http://absoluteshakespeare.com/guides/summaries/othello/othello_summary.htm
2. Breban, S., Vassileva, J.: Long-term coalitions for the electronic marketplace. In: Proceedings of the E-Commerce Applications Workshop, Canadian AI Conference (2001)
3. Brin, S., Page, L.: The anatomy of a large-scale hypertextual web search engine. In: Seventh International World-Wide Web Conference (WWW 1998) (1998). <http://ilpubs.stanford.edu:8090/361/>
4. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n -person games. *AI* **77**, 321–357 (1995)
5. Dunne, P.E., Wooldridge, M.: Complexity of abstract argumentation. In: I. Rahwan, G. Simari (eds.) *Argumentation in Artificial Intelligence*. *SV* (2009)
6. Erriquez, E., Hoek, W.V.D., Wooldridge, M.: An abstract framework for reasoning about trust. In: Proceedings of the 14th International Workshop on Trust in Agent Societies (2011)
7. Falcone, R., Castelfranchi, C.: Transitivity in trust. A discussed property. In: WOA (2010)
8. Felsenthal, D.S., Machover, M.: *The Measurement of Voting Power*. Edward Elgar: Cheltenham, UK (1998)

9. Griffiths, N., Luck, M.: Coalition formation through motivation and trust. In: Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems (2003)
10. Lei, G., Xiaolin, W., Guangzhou, Z.: Trust-based optimal workplace coalition generation. In: Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on, pp. 1–4 (2009). DOI 10.1109/ICIECS.2009.5366705
11. Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: Bring-ing order to the web. Technical Report 1999-66, Stanford InfoLab (1999). <http://ilpubs.stanford.edu:8090/422/.Previousnumber=SIDL-WP-1999-0120>
12. Qing-hua, Z., Chong-jun, W., Jun-yuan, X.: Core: A trust model for agent coalition formation. In: Natural Computation, 2009. ICNC '09. Fifth International Conference on, vol. 5, pp. 541–545 (2009)
13. Rahwan, I., Simari, G.R. (eds.): Argumentation in Artificial Intelligence. SV (2009)
14. Rahwan, T.: Algorithms for coalition formation in multi-agent systems. Ph.D. thesis, University of Southampton, UK (2007)
15. Sabater, J., Sierra, C.: Reputation and social network analysis in multi-agent systems. In: AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems, pp. 475–482. ACM, New York, NY, USA (2002). DOI <http://doi.acm.org/10.1145/544741.544854>
16. Sandholm, T., Larson, K., Andersson, M., Shehory, O., Tohmé, F.: Coalition structure generation with worst case guarantees. *AI* **111**(1–2), 209–238 (1999)
17. Wetzler, I.M., Zeelenberg, M., Pieters, R.: Never eat in that restaurant, i did!: Exploring why people engage in negative word-of-mouth communication. *Psychology and Marketing* **24**(8), 661–680 (2007)
18. Wooldridge, M.: An Introduction to Multiagent Systems, 2. edn. Wiley (2009)