

Applied Computational Trust in Utilities Management: a Case Study on The Town Council of Cava de' Tirreni.

Pierpaolo Dondio¹, Edmondo Manzo² and Stephen Barrett¹

¹ Trinity College Dublin, School of Computer Science and Statistics,
Westland Row 2.1 Dublin, Ireland {dondiop,stephen.barrett}@cs.tcd.ie,
WWW home page: <http://www.cs.tcd.ie/~dondiop>

² Consultant (until 31.12.2006), City Council of Cava de' Tirreni (SA),
Italy
emanzo@inwind.it

WWW home page: <http://www.comune.cava-de-tirreni.sa.it>

Abstract. This paper describes an application of computational trust techniques to enhance the water-supply service information system of the Town of Cava de' Tirreni, Italy. The study case covers a population of almost 52 000 people and about 23 000 consumption contracts. The Town Council is responsible for the water supply service and its billing process. A correct billing process requires gathering water consumption data of each citizen, task that is partially not controllable by Town Council personnel and therefore cannot be accomplished properly without the citizen's cooperation. Bad or malicious data are potentially harmful for both parties. The aim of this experimentation is to exploit computational trust techniques to better manage the process of validation of the received data. Computational Trust added value is represented by its autonomic implementation and by its in-depth consideration of social and environmental variables that go beyond simple data validation. The evaluation section, covering 6 years of data, will present the encouraging results obtained.

1 Introduction

This paper proposes an application of computational trust techniques for better managing specific tasks in a large public-sector information system. The study case described here refers to the water supply service in the Town of Cava de' Tirreni, in southern Italy. The Town Council is the only responsible for the water supply service and the billing process. The billing process is computed on the basis of the actual amount of water used by a single user, traced by a counter installed in the user's property. In order to accomplish its tasks, the Town Council needs the collaboration of the citizen.

The aim of this experimentation is to exploit computational trust techniques to better manage the process of analysing the validity and trustworthiness of the data that the Town Council received from its personnel or from citizens, that can be both affected by human mistakes, maliciousness, lack of experience, failures. An error in the data implies an error on the billing process, damaging the Town Council or the citizen, increasing the claims, delaying the payment process, delaying the Town Council procedures. The water service legislation has recently been changed and the modifications introduced make the present case study more meaningful. Under the new legislation, described in section 2, the right computation of each citizen consumption data became the critical process of the water service management.

Town Council workers are used to gathering consumption data over council territory. However, workers are actually limited in number and not exclusively assigned to this specific task; usually a massive reading of data over the territory requires the employment of fixed term workers and can be accomplished about once a year. External causes (like the lack of accessibility to the property, or the absence of the citizen) can limit the number of data gathered using this channel. Thus, the collaboration of the citizen in such a process is critical.

The analysis of the interactions involved in the service suggested to the authors an enhancement of the process using computational trust. The data transmitted by consumers are partially out of control of the Town Council and malicious values can damage the environment. This satisfies the main requirement for trust to be meaningful: there is a lack of monitoring of an action that may potentially hurt the parties involved or facilitate one party to the detriment of the other.

The authors of this paper chose to use computational trust to enhance the data validation process. Moreover, the automatic way in which computational trust techniques can be embedded in our information system is seen as a helpful added value.

We also expect what S. Marsh called the reduction of complexity [1], obtained by considering only trustworthy possibilities among a set of possible actions.

Finally, the accomplishment of this work is also motivated by the lack of applications of computational trust in information system management and, according to many authors [2], a gap between recommendation/reputation systems and other applications..

The aim of such an experimentation is now described. We seek to investigate the use of computational trust techniques in a new application field, where the same task is accomplished by simple data validation procedures, expecting an improvement. The data available in our study case cover 6 years of service, three of which managed with traditional data validation and three with trust-enhanced techniques. This will allow us to compare the two approaches. Computational trust techniques are expected to obtain more accurate predictions and, as a long-term effect, an increasing trustworthiness of the service perceived by the citizens.

Our application of computational trust techniques to information system management is a first effort to go beyond the data validation one. Many public bodies that manage a scenario similar to ours adopt data validation techniques to test the validity of data. These techniques in general do not go beyond a comparison among data values, possibly using statistical analysis.

On the contrary, we assume that our situation is not a mere problem of data validation (*are these data valid?*), but it should also be treated as a problem of trust

(*can I trust these data?*). Stating this means a deeper and broader understanding of the situation. Trust is not only concerned with the validity of the data, its meaning referring to the entity that issued those data and the context in which they were issued, considering past history as well as social interactions, domain-specific expertise, present evidences and other soft factors. In our scenario this means considering if the citizen is a regular payer or if the worker is competent. Thus, the techniques applied in this work range from simple data validation to more trust-specific techniques.

A trust analysis contains more information than simple data validation. For example, data that clearly result out of standard are for data validation techniques simply an error. From a trust perspective, if the data were issued by trustworthy entities, the error could be a mechanical failure or a mistake rather than a malicious attempt, situation more likely when the entity is not trustworthy. This extra information could possibly turn into different actions to correct the anomaly: simply contact the trustworthy citizen for another data reading or warn him of a possible failure; send a Council worker or a legal notice to the untrustworthy user.

By using trust-based techniques we seek to go beyond data validation for two reasons: a better accuracy on the predictions and more motivated and comprehensible reasons to justify the decision-making process. Data validation cannot grant trust, it only detects data anomalies based exclusively on the data value. Trust-added value begins exactly when data validation ends: given that certain data have some anomalies, can I trust them or not? If I grant trust I accept the suspected data anomalies. If the decision was correct – the ultimate goal of trust-based technique - I would avoid additional checking workload without harmful consequences.

Moreover, trust techniques are proactive while data validation is concerned with data only once they have arrived. For example, the generic trust mechanism *pluralism* states that data confirmed by multiple sources, better if independent, are more trustworthy. Thus, this implies to re-organize the process of data collection standing by this principle, in our case the rotation of the personnel over different areas of the Town Council during the data gathering process.

The paper is organized as follows: section 2 describes our scenario and the processes involved in the water supply service, section 3 describes related works that encompass similar approaches in the area of utilities management and computational trust references, section 4 describes the trust techniques that can suit our problem domain, section 5 describes the mechanisms that have been put in place to enhance the data validation procedure, section 6 describes our evaluation and finally we describes our conclusions and possible future extension of the project.

2. Water Supply Management System in Cava de' Tirreni

This section describes the experimentation scenario, explaining the water-supply service processes, actors involved and the dimensions of our study case. The Italian water supply service is generally managed by each Town Council on the territory. In this work we performed our evaluation in collaboration with the Town Council of Cava de' Tirreni, that entirely manages the water-supply service on its territory.

Cava de' Tirreni is a town of about 52 000 people, situated in southern Italy, 50 km south of Naples and 9 km north of Salerno, as shown in fig. 1. Some characteristics of the process are relevant for our experimentation.

The Town has a territory of 36,46 km², its population distributed in 19 different districts and a main centre. Water supply service operations are carried on dividing the territory in 30 main areas. The number of contracts up to 31/12/2006 is 24 357: 20658 private and 3699 business. All the customer service operations are carried out by a dedicate office in the Town Council with a personnel of 4 people. The Town Council has also 5 permanent workers, dedicated to counters installation, maintenance and collecting consumption data. In peak periods it also hires 3 or 4 temporary workers for a period of 6 months.



Fig. 1. Campania Region and Cava de' Tirreni.

The water supply service is regulated by a national legislation that changed in 2002. The change, concerning the billing procedure, is very relevant for our scenario. Until 2002, the citizen had to pay his consumption of water according to a fixed minimum per annum, established on the contract. This minimum consumption was paid in any case, regardless of its actual utilization, and only the exceeding part of water was paid according to its actual amount. The minimum consumption was a quite accurate estimation of the expected consumption, and thus the exceeding part had usually a minor impact on the total income.

Between 01/01/2002 and 31/12/2004 the minimum consumption for domestic use was progressively reduced and, since 1st January 2005, this quota has been removed as the new legislation prescribes. The invoice sent to users since 2005 is entirely dependent on the actual quantity of water used. The process of collecting water consumption data becomes essential and, if not properly carried out, much more harmful for the parties involved than before.

The aim of the Town Council Office is to send invoices as accurate as possible, i.e. to collect the exact amount of water consumption for the highest number of users. A lesser amount in the invoice results in damages the Council, generally not because this income goes wasted (it is almost always gathered when the data are collected properly), but because this mistake affects the correctness of the Town Council balance. A higher amount damages the consumer and at the same time the Town Council itself, because the invoice is usually not paid for and claimed. The user will refer to the Council Customer Care, delaying the payment process, triggering extra control procedures done mainly by a Town Council worker and forcing the office to re-calculate some invoices.

The amount of each invoice is established by the data received by the Council or, if not available, on a statistically computed value. The data collection process is now analysed in detail. The water consumption data are traced by counters installed in every user's property. By law, once a year all the citizens' counters must be read by personnel of the Council. This procedure is called the "*massive reading of counter*", and it is carried out by dividing the Council territory in 30 areas and assigning different areas to permanent or temporary workers. Note that this procedure does not guarantee the complete collection of the data required, since some users may be temporarily absent or the property inaccessible. For each user there are usually multiple data collected in different periods over the years.

During the year data are collected by relying on the collaboration of the users and, in special case (verifications, suspected failures), by resorting to Town Council workers. In order to compute the amount of an invoice, all the consumption data received by the various channel are considered. In 2006 the total number of readings used to compute invoices was 26 192, 57% of which collected during the period of "massive reading" by permanent and temporary workers, 19% by the permanent term workers outside that period and 24% by the citizens. It is clear that users play an important role in the process: they send their consumption data - and that is when malicious users can transmit erroneous data or data in contrast with the reading done by operators.

3. Related Works

This section covers three aspects: computational trust techniques definition, current related applications of computational trust, how other public authorities or utilities manage a process analogous to ours.

3.1 Computational Trust techniques

Computational trust seeks to exploit the human notion of Trust into the digital world. A largely accepted definition of Trust in mathematical terms is the classical definition of Gambetta [3], where Trust is a particular level of the subjective probability with which an agent assesses that another agent will perform a particular action, both before he can monitor such action and in a context in which it affects his own action. Gambetta's definition influences the quantitative representation of trust as a probability, but, as noticed by Castelfranchi and Falcone in [4], it doesn't take into consideration mental processes and mechanisms involved in the quantification of such a probability that represents the ingredients of a trust-based decision. In our paper we underline the need to give clear justifications for a trust-based decision by making explicit motivations and listing the evidence behind a decision.

Fig. 2 represents the elements of a computational trust solution. A trust-based decision in a specific domain is a multi-stage process: the selection of the required input data (trust evidences) by mean of a notion of trust (trust model); the processing of the inputs producing a trust value, and the actual decision considering computed outputs and exogenous factors, like disposition or risk assessments. As a dedicated trust

infrastructure (PKI, Recommendations) is not part of the system considered in this paper (and its utility should be investigated), evidences must be directly collected and exploited selecting appropriate domain elements and dynamics. We apply evidence-based trust where the trust computation is supported entirely by elements of the application. One of the authors of this work performed previous application-contained trust computation in the Wikipedia project [5].

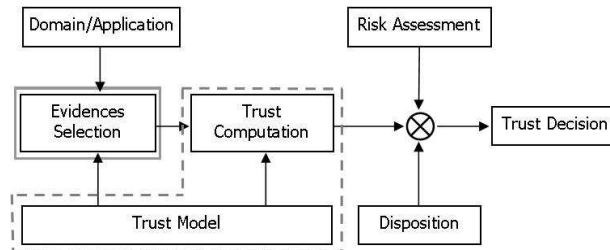


Fig. 2 A Computational Trust Solution

Computational trust identifies trust-based techniques that humans adopt in the decisions whenever trust or distrust is involved. These techniques are taken from interdisciplinary study of Trust, encompassing social science, psychology, cognitive approaches. In section 4 we will describe how we fit some of this mechanisms to our specific domain; here we briefly review the general patterns of these mechanisms referring to some researches in the area that used such mechanisms. The aim of this section is to give a reference to the mechanisms we implemented in our work.

The importance of memory and past history has been largely recognized in many trust models. For instance, the *trustcomp* community defined trust as “a non enforceable prediction of the future behaviour of an entity based on past evidence” [6]. In the *past-evidence paradigm*, implemented for example in the Secure trust engine [7], the trustor collects outcomes of past interactions to compute a trust value for the trustee. Past evidences should be collected in significant numbers, the external conditions should remain stable for the evidences to have sense. The entity with which we are interacting must be recognized in order to bind entity and outcomes. Outcomes (or trust values) can be received and shared with other entities in order to exploit third party experience to assess trust. The sharing of information is related to the Recommendation system, largely studied in centralized and decentralized environment. Related problems studied encompass privacy protection and the identification of malicious users. By exploiting the transitivity of trust values, a Social Network [8] can be constructed and trust values propagated through the links of the network.

A significant approach to trust is represented by the *probability-based* approach. Probability-based trust performs trust calculation by applying the theory of probability over relevant sets of data. In general, the underlying idea is that trust is the probability that the trustee will fulfil the expectation of the trustor. The method required that a sufficient set of triples (trustee, action, outcome) are collected in order to estimate the probability associated with a specific couple entity-action. Examples include Despotovic [9], who used the maximum likelihood theorem to predict good outcomes in a p2p recommendation system for ebusiness, and Wang [10], who applied the bayesian probability in eBay.

Similarity between entities is another mechanism to propagate trust values. Similarity as a source of trust was studied by Ziegler and Golbeck in [11]. In [12], Castelfranchi and Falcone proposed a computational model of their cognitive approach to trust, encompassing similarity in the sources of the mental believes composing trust. *Categorization* was indicated as another mechanism, where trust values are propagated among categories of entities. Their work consider also the concept of the *pluralism* and *multiplicity of sources* to enforce the trustworthiness, especially when many independent sources confirm the same data in different situations. Finally, risk management and assessment is often coupled with trust analysis (see Secure [7]), since the action of trust towards an entity could be not convenient from a risk point of view.

3.2 Computational Trust and Security application and Consumption Data collection in Utilities management.

Many formal trust models have been defined since 1996, when Stephen Marsh proposed the first model of computational trust. Despite these well-established researches, according to Seigneur [2], a “big effort has been put in defining a formal model of trust, but there are still many applications that needs to be considered and many evidences that need to be identified and processed”.

Computational trust methods have gained successful applications in e-commerce scenarios like amazon.com, online auctions systems like ebay, spam email filtering, recommendations and rating systems like epinions.com, movietrust.com (in the social network variant), online forums.

In business, computational trust has a lesser range of application as compared to parallel researches of risk management, security and privacy. The authors find it difficult to identify applications similar to the one presented in this paper, especially due to the peculiar field of application. While reputation-based and recommendation techniques seem to have several applications, the application of trust in business is mainly security-oriented: digital signature protection, privacy management, clustering of customers and sellers [13]. The latter approach, involving collaborative filtering, fuzzy logic-based classification is relevant to some classification models used in this paper. In general, computational trust techniques are not still exploited in the management of information systems specifically for an administrative task like ours. In this system, information is inserted relying on users' cooperation, and lack of control and malicious actions are possible. Data have often been kept safe using key infrastructures, but they do not guarantee their trustworthiness.

Many public Utilities manage a scenario similar to ours adopting only data validation. In the Province of Salerno many public companies are managing water supply services for cities up to 20.000 users [14]. Our survey of the used methods (limited to the Italian territory) showed how techniques beyond data validation are not implemented. The techniques adopted by the institution we analysed can be summarized in a common scheme. The consumption data received, collected by a wide type of channel, are rejected by the system if :

- a) in the consumptions database there are previous data greater than the last data consumption (situation technically impossible)
- b) the consumption database contains data equal to the previous one

c) the reading data are lower than those already involved in an invoice process.

If the data do not belong to any of the above situations they will be saved in the database, otherwise they will be rejected. The procedure guarantees data consistency and prevents only a mismatching event during the invoice calculation process, since the data rejected will not be involved in any invoice calculation. No other information except the data value is considered.

The same system is applied to other utilities management. Enel [15], the national Italian supplier of energy, based its billing on the same process.

In 1998 the Italian Ministry of Treasure established the *redditometro* (“income evaluation meter”), [16] a mechanism to identify tax returns with anomalous values. On a larger scale, this example is conceptually similar to ours: tax return is done by citizens and it is partially out of control by the state, since controls are obviously not possible on every single citizen. The mechanism works using the presumptive value that every citizen should pay their taxes according to their category. Each category of workers, assets or properties has an estimated expected revenue attached to it, computed performing average among categories of workers, the so called *ministerial study of professional categories*. The value of a tax return should not exceed a certain threshold for the expected values. If it does, more controls should be done. The system reflects the idea of having a predetermined amount, similar to the water supply system preceding 2002. Anyway, clear differences make our scenario more complex: the need for accurate values (expected values are not enough), the possibility of claims by the citizen in the short time, yearly balance requirements.

4. Fitting Computational Trust Techniques

In this paragraph we underline the elements of the process that we found useful for the application of the trust mechanism.

Categorization of users. The Town Council contracts are divided into two general categories: private and business. For each category there are some sub-categories that better specify the water utilization. Private sub-categories are defined as follows: residential home, secondary home, private garage, loft etc. If a house is not the users’ habitual residence (it is the case of a holiday house, for instance), then it is classified as a secondary residence; commercial or business sub-categories differ according to water utilization: restaurants, hotels, hospitals, hospices, laundries, industries, agriculture, cinemas, shops, etc. The categorization of the users will sustain the application of *categorization* and *similarity* trust mechanisms. The category specified by the contract is not merely a classification of the users but it implies the utilization of the water for the correct and declared scope.

Sources of consumption data. We can identify three channels of consumption data acquisition. The first channel (“O”) is represented by the data collected by permanent Council workers while carrying out their normal tasks. The second channel (“L”) is represented by the data acquired by the temporary (L1) and permanent (L2) workers during the “massive reading” period. The third channel (“A”) is data collected by the users. Data are collected dividing the council territory in 30 areas. The division is made to better manage the workers over the Council territory and to enhance the prediction of the reading, if necessary. In fact, in this last case, the workers collect the reading

data at least once a year in a default period in the same area. The classification of data sources allows binding past data analysis to its relative source of acquisition (past evidence mechanism), while the possibility of having different types of sources over the different areas sustains the use of *multiple sources* mechanisms.

Past history (consumption data). Each user's past data are recorded. Past invoices are recorded, along with all the data consumption readings used to compute each invoice. These historical data are the input required by statistical trust mechanism. We shall consider the number of data acquired, but also the number of data accepted and rejected by the system for each customer, from which we can deduct the quality of the users' data. The minimum consumption quota, that in the year before 2002 was the key datum used for computing the invoice, is also taken into account as a statistical reference.

Family Composition. Trust mechanisms based on past data or past interactions rely on the stability of the entity under analysis: a change in entity's properties can invalidate past data and trust computation. The number of people in a family is an essential information to estimate the consumption data. Components' number may vary during the year and the contract may not have been updated (or there is no possibility to change it). This is why we linked our information system with the council registry database, to avoid that a change in the family composition could invalidate past data and erroneously alter the reputation of a user.

Payments Status and Invoices Status. Past payment dates and status are an important piece of information. We can understand if the user usually pays regularly, with delays or if he was forced to pay. There are five levels of notice that the Town Council can issues, from the simple warning to various legal notices and contract suspension. We considered also if the citizen received some reimbursement due to an erroneously overcharged invoice. These data represent the past interaction between the citizen and the Town Council and an important source of mutual trust evaluation.

Failures: The Town Council is responsible for the maintenance of the public pipeline network; each possible failure in the public network is monitored and controlled by the Town Council, while the end-users' pipes, from the counter to the domestic network, are the users' responsibility. Only a Town Council operator can change a counter; thus it is possible to trace the counter installation and reparation done by a specific operator. Using this information, if a counter has just been replaced, it is unlikely that the transmitted data are erroneous because multiple failures are unlikely. In this case trust is coupled with event probability.

5. A trust-based enhanced system

In order to provide a computational trust solution, we had to identify which elements of the domain are trust evidence, as we did in section 4, define one or more trust metrics, i.e. how to evaluate the trust evidence gathered to produce a trust value TV and decide to trust or not according to the trust value and the set of evidence.

Table 1 shows the elements used in our trust metric. The relative importance of each element in the final aggregated trust value is specified. Our trust metric is case-

sensitive, and it computes trust values usually in two stages. Stage 2 is considered only if the resulting trust value after stage 1 is not enough to accept data.

Table 1. Trust factors table. The impact of each factor is shown in the 3 general cases, with the stage of computation (in brackets).

Trust Factor	Case a	Case b		Case C	
		Stage 1	Stage 2	Stage 1	Stage 2
Sub-category		High		Medium	
Past Consumption		Medium		High	
Data					
Past Payments	Low		Medium		Medium
Number of Readings		Low		Medium	
Channel of acquisition	Low	Medium			High
Source Rotation		Low		Low	
Minimum Quota				Medium	
User's past data rejection rate					Medium
Counter Maintenance	Low		Low		Medium
Family Status				High	

Besides the trust value for a specific data, our system keeps track also of a global trust value, not referring to specific consumption data but to the users, i.e. a value that summarizes all users' previous data validation sessions.

The trust-based system can be used in many ways. A trust value can be produced in an autonomic way, represented by a numeric value in $[0..1]$. However, this is a limited representation that may hide (or lose) details. That is why the system can list all the reasons behind a computation, showing all the trust evidences used and their individual value for the user under analysis.

All the information is also available to the help-desk operator through a new user-end interface, in order to facilitate a human-based analysis. The operator can query the system and handle special and fuzzy cases, usually in front of the user in the Help-Desk office, using soft skills and direct knowledge of the user.

The consumption data for each contract are saved in a *consumption database* that we enhanced to support our trust information. The database shows, for each consumption datum, the date of acquisition, the counter's value, if the data are involved in any invoice, the acquisition's channel, the worker who collected them and their state (active/not active). Active data are valid and they contribute to invoice calculation while non active data are data rejected by the invoice calculation process.

For the Town Council operator, two data panels (reading data panel and user panel) have been added to facilitate the analysis of data and support the decision making process. The panels contain information on the family composition as well, taken from the civil registry database that we linked. The examination is helped by the use of the *consumption data panel* application: it allows the operator to divide the cases of evaluation and it gives each contract a complete overview of water consumption. The panel presents three different filters: a) invalid consumption; b) zero consumption; c) greater or lower than expected. These are the three general cases, that are treated differently: case a is a clear situation that can be faced with simple data validation, while case b and c are complex cases requiring a trust-based analysis. The aim is to

validate consumption data, that should be the exact amount, or to predict the water consumption. The evaluation system of the reading data is shown in figure 4 .

Contratto	Contratto	M.C.	Letta Fatturata	Ultima Lettura	Differenza	MC	MC	Consumo	Tipo Lettura	Tipologia Contratto	Destinazione d'uso
Interno			Data	Data	Giorno	Giorno	Attuale				
15475	15475	543	14/05/1996	442	05/05/2001	5104	1817	1,48	2.898,16	Standard Diverso	
75295	74041	720	31/12/2005	6072	13/09/2006	13696	256	1,97	504,99	Standard Diverso	Industria
71831	71397	180	31/12/1999	170	31/12/2004	4707	1527	0,49	940,99	Standard Domestico	Domicilio famiglia
91495	81495	1930	07/12/2009	10201	01/12/2003	14925	724	4,83	5.250,41	Automatica Palmare	Casale, casa di riposo
96414	96414	1930	22/11/1999	1	24/05/2001	3449	549	4,83	2.707,40	Standard Diverso	Ristorante, pizzeria, rosticc
11540	11540	2160	31/08/1996	49225	07/01/1997	52533	462	5,92	2.339,98	Standard Diverso	
59173	59173	10930	31/12/2005	28339	09/03/2006	31334	68	29,59	2.012,05	Automatica Palmare	Collegio, convitto
15238	15238	270	31/01/1994	2216	16/03/2001	4702	2573	0,74	1.903,32	Standard Diverso	Domicilio famiglia
4259	4259	109	31/12/1994	1143	29/05/2001	3694	2341	0,30	602,68	Standard Domestico	
4656	4656	216	07/03/1997	1	18/09/1997	2100	103	0,59	60,95	Standard Domestico	Domicilio famiglia
9528	9528	720	20/12/1995	520	30/11/1997	2465	711	1,97	1.402,52	Standard Diverso	Hotel, pensione, motel
16067	16067	108	31/03/1994	1160	17/09/1997	3002	1266	0,30	374,60	Standard Domestico	
53491	53491	270	10/09/1995	362	25/07/2001	2150	2176	0,74	1.609,64	Standard Domestico	Domicilio famiglia
17943	17943	540	11/05/1994	11745	31/12/1994	13521	234	1,48	369,16	Standard Diverso	Casale, bar, pasticceria, gel
16275	16275	72	22/12/1994	933	28/04/2001	2523	2319	0,20	457,45	Standard Diverso	Sala da ballo
70773	70773	390	15/12/1997	3448	10/04/2001	5090	1212	0,99	1.195,40	Standard Diverso	Azienda commerciale
52356	52356	378	09/11/1997	714	31/12/2001	2263	1514	1,04	1.587,52	Standard Domestico	Domicilio famiglia
76564	70202-1	7200	31/12/2005	60158	25/01/2006	61955	25	19,73	493,15	Manuale Operatore	Ufficio pubblico
15596	15596	2160	31/12/2000	3662	07/09/2003	5001	219	5,92	1.296,00	Standard Diverso	Industria
1589	1589	180	17/05/1995	3645	27/09/1999	5163	1594	0,49	796,08	Standard Diverso	
5348	5348	288	07/07/1997	965	28/10/2002	2469	1930	0,79	1.529,95	Standard Domestico	
9803	9803	108	15/12/1994	99	31/12/2001	1600	2573	0,30	781,33	Manuale Autolettura	Domicilio famiglia
8548	8548	216	03/01/1999	1	07/01/2003	1487	1465	0,59	866,96	Standard Domestico	Domicilio famiglia
56993	54993	324	03/06/1995	469	31/12/1999	1953	1672	0,88	1.649,15	Standard Domestico	Domicilio famiglia
16930	16930	216	31/12/1995	4184	31/12/2002	5831	1461	0,59	864,59	Standard Domestico	Residenza secondaria
53122	53122	1090	31/12/2005	25870	09/06/2006	27319	160	4,83	789,04	Standard Diverso	Oratorio, casa di riposo
70350	70350	199	10/11/1997	521	07/12/2001	1949	1488	0,49	733,81	Standard Domestico	Domicilio famiglia
16383	16383	180	10/01/1997	1160	17/03/2000	2569	1162	0,49	573,04	Standard Domestico	Ufficio pubblico
15916	15916	360	03/03/1994	107	16/09/1996	2450	928	0,98	915,26	Standard Diverso	Domicilio famiglia
17331	17331	180	02/01/1995	369	19/11/1999	1729	1782	0,49	879,79	Standard Domestico	Domicilio famiglia
72444	52448-1	1090	31/12/2005	18964	02/07/2006	17260	168	2,96	541,48	Standard Diverso	Industria
866	866	198	04/02/1994	620	26/09/1997	1946	1330	0,49	655,95	Standard Domestico	
3552	3552	540	15/03/1996	213	31/07/1998	1537	138	1,48	204,16	Standard Diverso	Domicilio famiglia
59515	50515	216	20/01/1994	434	17/02/1998	1775	1489	0,59	691,16	Standard Domestico	Domicilio famiglia
15022	15022	108	14/03/1994	2613	31/12/2002	3623	3214	0,30	950,99	Standard Domestico	Domicilio famiglia
94132	77493	5400	31/12/2005	37105	27/03/2006	38396	86	14,79	1.272,33	Automatica Palmare	Clinica, casa di cura, Anali
17321	17321	288	23/09/1995	1721	30/05/2000	3015	646	0,79	699,72	Standard Domestico	Domicilio famiglia
9764	9764	108	19/01/1996	1	31/12/2001	1296	2173	0,30	842,97	Standard Domestico	Domicilio famiglia
17693	17693	540	13/03/1994	3933	04/01/1996	4764	571	1,48	944,77	Standard Diverso	Industria
2114	2114	1620	03/08/1994	17155	01/05/1995	18381	271	4,44	1.202,79	Standard Diverso	

Fig. 3. Consumption data panel

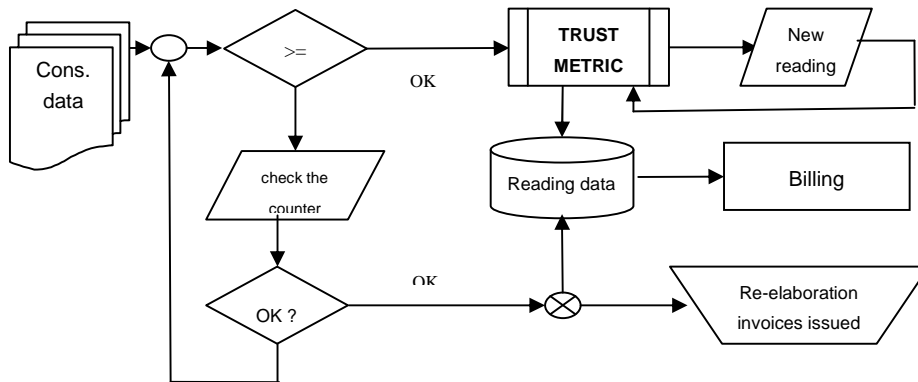


Fig. 4. Evaluation system

Case a: last consumption data are lower than the previous. In this case data are invalid and rejected (total consumption cannot obviously decrease). Trust is not applied to this situation, but only mere data comparison. Invalid data may be a consequence of the citizen makes confusion with other counters like gas. New reading data are collected by a permanent worker (only permanent workers can check and, if needed, replace the counter). These new data are compared by the rejected one and, if the situation is confirmed and the counter is working properly, this means that past data and correspondent invoices were overestimated. The operator re-calculates the last one

and, if necessary, also the previous water consumption invoices, issuing a reimbursement note in favour of the citizen.

The consumption data accepted by the process (case b and c) are subject to the “data evaluation system” based on computational trust techniques.

Case b: evaluation of reading data that are equal to the previous. In this case, the data consumption was less than 1 m^3 , the resolution of the counter. A trust metric is applied. As shown in table 1, the metric starts considering the sub-category of the contract (high impact): only some specific categories, like garages or stores, have expected consumption data that can be less than 1 m^3 per year. If the contract doesn't belong to one of these categories its trustworthiness will decrease drastically. Anyway, past user data are also considered by the metric, with a strong impact if they have been done by Town Council Operator (channel O and L) and a lesser impact if collected by the user. If several past data by channel O or L confirm a consumption close to zero, data are accepted. In case of low trust value after stage 1, the hypothesis of malicious data is usually discarded since, according to our domain expert, a malicious user usually declares the same level of consumption, in order not to raise suspicions. The counter is probably broken and halted, or the reading data are erroneous. In this case, with a low value after stage 1, an operator is sent to check the counter. The trust metric considers now the past payment status of the citizen to determine the next user trust value. If the citizen was a regular payer, the data are considered true and the possibility of a mistake or failure higher. The user trust value is not lessened. If the citizen had a bad payment history or a recently replaced counter, user's trust value is lessened and, if the user's trust value reaches a lower threshold, the Town council may decide to send a legal notice instead of an operator control, in many case useless.

Case c: evaluation of reading data that are greater than the previous ones. This is the core of our experimentation, the most frequent and complex case.

The data evaluation module verifies the consumption data, like in case b. Sub-cases are possible:

1) Data are as expected (i.e. trustworthy), so the reading data successfully pass the evaluation. In this case the trust metric computation considers if the data are compatible with the expected data, computed on statistical past consumption data used in past invoices (thus accepted). The trust metric considers the number of data available (more data, higher trust value), the period covered (short period, lower trust value), the channel of acquisition (more acquisitions by permanent worker, higher trust value). If data are not enough, the old minimum consumption quota is considered, but the resulting trust value is lower than the one computed with user-related consumption data. The sub-category is considered as well, enforcing or decreasing the trust value with a medium impact. According to this consideration, data may have enough trust value to be accepted, increasing the user's trust value.

2) greater/lower than expected. The data consumption is not applicable from a data validation point of view. We need a more sophisticated trust-based analysis to understand its meaning.

Data are greater than expected. Trust metric starts evaluating how much the data are above the old minimum consumption quota. If the data are compatible, the metric considers the sub-category. Even if the data are correct, the consumption may not be compatible for the category. For example, this is the case of a contract declared for a specific use (such as a shop) and actually applied to a factory. This affects the general user trust value, since the user is not using the water as his contract allows. This reason,

when the user's trust values are decreased several times for this reason, can lead to an inspection by the Town Council operator and the user can be forced to change contract and pay a penalty fee.

After stage 1, the trust metric considers now the payment status of the user and the channel of acquisition. If the data were taken by a the citizen and the citizen is a regular payer, the user's trust value is not changed and usually the system suggests to contact the user for another data transmission, to double check it. The user is trustworthy; there is no need to send an operator. If everything is confirmed, the data are accepted and user's trust value is unchanged, since the situation is considered anomalous. In fact, only case a) can increase the user's trust value. The trustworthy user may have a leakage in his pipe network he did not realize and by contacting him the Council increases its Service Quality in the respect of its citizens. If the user is not a good payer - for example the user doesn't pay the invoices because he thinks the consumption invoiced is not right -, a Town Council Operator is sent to check the reading.

Data are lower than expected. If the data are lower than expected, the subcategory and the channel of acquisition play an important role in the trust metric. The other trust factors (table 1) are similar to the previous case and the procedure followed is the same. The difference is that this case has a worse impact on the user trust value than in previous cases. In case the counter is checked by a Town Council operator and the low consumption is not confirmed, the user's trust value decreases. Note that, from the Council point of view, the previous case is more critical, since the Council is a public body and thus considers better to underestimate an invoice than to make the user pay more than the right amount.

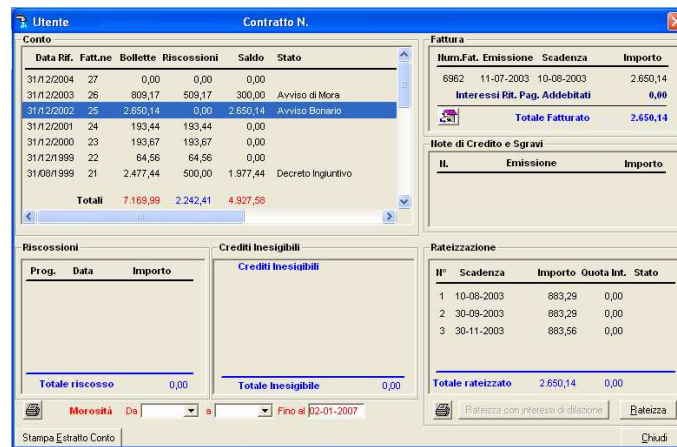


Fig. 5. Consumer overview panel

Finally, for the cases in which a trust metric is used (b,c), the trust value decreases gradually in time until there is a Town Council Operator intervention. For instance, even if the data successfully passed the validation, if the last reading by an operator was two years ago, all the trust values are reduced. When the trust value decreases under a certain threshold, a reading by a Town Council operator is needed to boost the trust value again. When this analysis is done interactively by a help desk operator, the

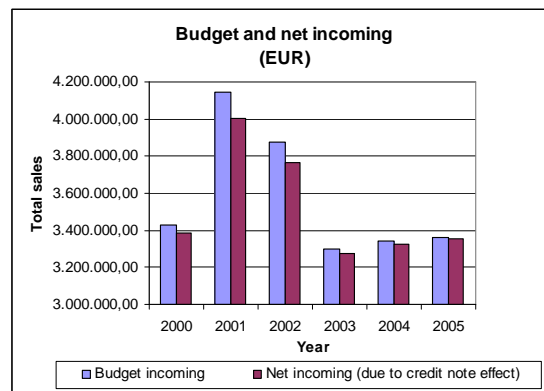
consumer overview panel permits to see the majority of the requested information. This panel (fig. 5) allows the operator to know the total invoice issued, the payment status, the credit notes issued, the irrevocable credits, the collected credits for each contract.

6. Evaluation

Our evaluation covers six years of data, from 2000 to 2005. Our enhanced mechanism became fully operative in 2003. Before applying computational trust techniques the Town Council Customer Care experienced (i) long queues at the help desk, (ii) frequent claims by consumers, that implied extra workload for Council desk operators, (iii) the claims were usually accepted by the Customer Care, meaning the issue of loads of reimbursement notes and deteriorating the level of trust of the citizen towards the Town Council. The introduction of computational trust techniques should give a more accurate estimation and thus reduce the claims. Our evaluation is based on the percentage of income estimated over the actual total income received by the Town Council. The difference between these two values is the total amount of invoices claimed by the citizen and reimbursed. If this percentage is low, this means that predictions were accurate and few invoices were claimed. The total amount of income is not a valid metric, since it could vary, and it is not a Town Council goal to increase the total income (a public bodies is non-profit) but to predict the right income. Our system, started in the year 2002, in few years provided the following advantages:

1. reduction of claims at the customer care desk;
2. reduction of the total amount of credit notes issued;
3. reduction of the difference between estimated and actual incoming

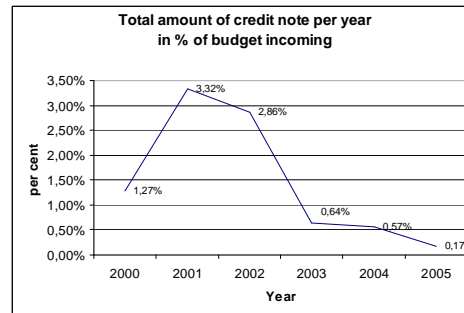
Graph 1. Actual and Estimated Incoming.



Graphs 1 and 2 display the result of our experimentation. Before proceeding to discuss the graphs it is advisable to underline that in the fiscal year 2000 most of the invoices were calculated according to the minimum quota consumption. In the year 2001 the Council introduced the “*massive reading of counters*” that became fully operative in the year 2002. With this new process a greater number of reading data was acquired than in the past. Consequently, the invoices issued for the fiscal year 2001 and 2002 contained also the extra consumption of the previous years (max 5 years before)

if the user exceeded the minimum quota. In fact, only when consumption data are available is it possible to calculate the exceeding part of the water. This explains the high income of the fiscal years 2001 and 2002. Since 2003 the income has been almost constant (some increments on the charges and a growth of total users explain the small increase), which by itself represents a good result: big variance in the amount means bad previous estimations, as is the case of the remarkable gap in year 2000, 2001 and 2002.

Graph 2 summarizes the reduction of the total amount of reimbursements. This is directly linked to the claims done by the citizens, and their number (since the amount of reimbursement has a low variance). Graph 2 shows the percentage of reimbursement over the total yearly amount. We can notice that in the fiscal year 2000 most invoices were based on fixed - minimum consumption and that year credit notes were issued for 1.27% of the total budget income. For the years 2001 and 2002 this percentage strongly increased while it began to decrease after the introduction of our enhanced system. In the year 2005 the total amount of credit notes issued is 0.17%, a clearly encouraging low value. We believe these results are not only due to a more precise system, but also to a better cooperation on the side of the users, even if quantitative metric for evaluating this cooperation is not yet implemented.



Graph 2. Percentage of income reimbursed over the actual total income.

7. Conclusions and Future Works

In this paper we described an application of computational trust techniques to enhance the water supply information system of Cava de' Tirreni, a middle size Town Council in Italy. We described the water consumption invoice system in detail, a process where citizens' collaboration plays a critical role. We showed how the process of water consumption validation, necessary for a correct invoice system, can be enhanced by introducing, along with usual data validation techniques, computational trust techniques. Using the latter approach data are accepted or rejected not only on the basis of their numerical values, but also considering the quality of the source by which the data have been collected, the reliability of the identity issuing the data and some statistical and environmental variables that may affect the data.

The experimentation has covered a period of observation of 6 years. Our results showed how the trust-enhanced predictions have achieved better results: the difference

between the estimated income, based on predictions, and the real income – after the necessary checks – has continuously decreased. This is reflected on the number of invoices claimed by the citizens and cancelled by the system. These results have brought many benefits for the Town Council, such as lesser delays in payments and a reduced workload for Council customer care operators. In our future developments, we will focus on a more detailed tracking of the process and we will seek to make the process more autonomic. We will also seek to apply the system on a larger scale, involving water supply companies that are managing several Town Councils services. In a larger scale experiment, several processes that in our experiment are still carried out manually because of the limited number of variables –such as the management of Town Council operators - will be regulated by autonomic trust procedures based on past history and recommendations.

8. References

- [1] Marsh, S. Formalizing Trust as a Computational Concept. PhD thesis, University of Stirling, D. of Mathematics 1994
- [2] Seigneur J.M. *Ambitrust? Immutable and Context Aware Trust Fusion*. Technical Report, Univ. of Geneva, 2006
- [3] D. Gambetta, *Trust: Making and Breaking Co-operative Relations* Basil Blackwell, Oxford 1988.
- [4] Castelfranchi, C., Falcone, R.. Trust is much more than subjective probability: Mental components and sources of trust. 32nd Hawaii International Conference on System Sciences, 2000.
- [5] P. Dondio et al. Extracting trust from domain analysis: a study on Wikipedia, IEEE ATC, Wuhan, China, 2006
- [6] Trustcomp group. Website www.trustcomp.org
- [7] V. Cahill et al., Using Trust for Secure Collaboration in Uncertain Environments. IEEE Pervasive Computing, Vol. 2, N. 3, July-September 2003
- [8] Golbeck, J., Hendler, J., and Parsia, B., *Trust Networks on the Semantic Web*. University of Maryland, College Park, USA, 2002
- [9] Z. Despotovic, Maximum Likelihood Estimation of Peers performance in a P2P network. CIA, Erfurt, Germany, 2006
- [10] Wuang Y. et al. Bayesian Network Based Trust Management, IEEE ATC, Wuhan, China, 2006
- [11] Castelfranchi, C., Falcone, R., Peluzzo, G. Trust in information sources as a source for trust: a fuzzy approach. AAMAS, Melbourne, Australia, 2003
- [12] C. Ziegler, Golbeck J. Investigating correlation between Trust and Similarity. Decision support system 2005.
- [13] Meier A., Werro N. Extending a WebShop with fuzzy classification model for online customer. IADIS eSociety Conference, Dublin, Ireland, 2006.
- [14] www.atosele.it Water Supply Service Consortium (ATO4) of the Province of Salerno, Italy
- [15] Enel, Ente Nazionale Energia Elettrica, www.enel.it
- [16] www.agenziaentrate.it, website of the Italian Ministry of Treasure