A Multi-Dimensional Trust Model for Heterogeneous Contract Observations

Steven Reece and **Stephen Roberts**
Department of Engineering Science
University of Oxford
Oxford, OX1 3PJ, UK.
{reece,sjrob}@robots.ox.ac.uk

Alex Rogers and **Nicholas R. Jennings**
Electronics and Computer Science
University of Southampton
Southampton, SO17 1BJ, UK.
{acr,nrj}@ecs.soton.ac.uk

Abstract

In this paper we develop a novel probabilistic model of computational trust that allows agents to exchange and combine reputation reports over heterogeneous, correlated multi-dimensional contracts. We consider the specific case of an agent attempting to procure a bundle of services that are subject to correlated quality of service failures (e.g. due to use of shared resources or infrastructure), and where the direct experience of other agents within the system consists of contracts over different combinations of these services. To this end, we present a formalism based on the Kalman filter that represents trust as a vector estimate of the probability that each service will be successfully delivered, and a covariance matrix that describes the uncertainty and correlations between these probabilities. We describe how the agents' direct experiences of contract outcomes can be represented and combined within this formalism, and we empirically demonstrate that our formalism provides significantly better trustworthiness estimates than the alternative of using separate single-dimensional trust models for each separate service (where information regarding the correlations between each estimate is lost).

Introduction

Computational models of trust have recently generated a great deal of research interest within the academic literature of multi-agent systems. Such models allow agents to select between various suppliers of services on the basis of their reliability or trustworthiness. To be effective, these models should allow agents (i) to estimate the trustworthiness of a supplier as they acquire direct experience, (ii) to express their uncertainty regarding this estimate, (iii) to exchange their estimates as reputation reports, and (iv) to filter and fuse these reputation reports with their own direct experience to yield more accurate estimates.

While much of the work within this area has used domain specific or *ad hoc* trust metrics (see Ramchurn, Hunyh, & Jennings (2004) for a review), a growing body of research shows that the desiderata described above may be achieved through grounding computational trust models in probability theory. Specifically, models of this form have been presented by a number of researchers, and typically they use a beta distribution to represent an agent's belief that a supplier will successfully fulfill a single-dimensional contract (Jøsang & Ismail 2002; Teacy *et al.* 2006).

In recent work we have extended such probabilistic trust models to consider cases in which a contract's success or failure is measured over multiple dimensions (Reece *et al.* 2007). These cases are common in real-world applications (e.g. within a supply chain where a contract specifies minimum timeliness, quality and quantity criteria), and in such cases, we would expect there to be correlations between the success or failure of each contract dimension (e.g. suppliers may trade-off failure in one dimension to achieve another; sacrificing quality or shipping a partial order to achieve a delivery deadline). In this context, we have shown that it is essential to explicitly consider these correlations if the expected utility of interacting with any particular supplier is to be accurately estimated. We have presented a formalism whereby the Dirichlet distribution (a natural multi-dimensional extension of the beta distribution) is used to represent an agent's correlated beliefs regarding these multiple contract dimensions, and we have described how these beliefs may be communicated between agents as reputation reports, and fused with an agent's own direct experience.

However, this formalism explicitly assumes that all agents observe and record contract outcomes over an identical set of dimensions (i.e. that the observations of contract outcomes that constitute each agent's direct experience are homogeneous). This limitation means that it can not be applied in the more general setting where correlations exist but observations of contract outcomes are heterogeneous. For example, consider the case of an agent attempting to negotiate with a supplier to procure a bundle of video, audio and data services in order to facilitate an interactive video conference. We would expect there to be correlations between the probabilities that each service will be successfully delivered (due to the fact that the services may share common resources or infrastructure such as communication networks, routers and servers), and thus, in order to estimate the expected utility of interacting with the supplier the agent must estimate these probabilities and correlations. In order to do so, the agent would benefit from combining its own limited direct experience with reputation reports from other agents. However, in this case, these other agents may only have experience of subsets of the entire bundle (i.e. one agent may only have used audio services provided by this supplier, while another may have experience of procuring both video and audio services but not data services). As yet, no principled computational trust model exists that allows an agent to combine these heterogeneous contract observations, while still maintaining information regarding their correlations.

To rectify this shortcoming, in this paper we develop just such a computational trust model. In doing so, we adopt a formalism common within the academic literature of target tracking and data fusion, and we use the Kalman filter to combine these heterogeneous contract observations. This approach is attractive since not only does it provide a solution to the problem at hand, but it also enables other results within the data fusion literature such as the gating of inconsistent estimates and the elimination of rumour propagation within decentralised information systems to be naturally incorporated into future computational trust models.

In more detail, we first show that in order to estimate the expected utility of a bundle of services, an agent must use a trust model that allows it to estimate (i) the probability that each service will be successfully delivered, and (ii) the correlations between these estimates. We then build on this model and make the following contributions:

- We develop a benchmark trust model for dealing with heterogeneous contract observations that uses separate single-dimensional trust models (specifically independent beta distributions) for each individual service within the bundle. This approach provides consistent estimates but does not represent correlations between the services.
- We describe a novel formalism that uses the Kalman filter to combine agents' heterogeneous contract observations while also explicitly representing the correlations between the services. We show how agents can calculate prior trust estimates and reputation reports from their own direct experience, and how these can be fused together to yield posterior trust estimates.
- We empirically demonstrate that by explicitly capturing the correlations between the services, our formalism based upon the Kalman filter yields far more precise estimates of the trustworthiness and expected utility compared to the alternative approach of using independent beta distributions. In our experiments the information content of estimates derived from the Kalman filter is typically three times that of estimates derived from the independent beta distributions.

The remainder of this paper is organized as follows: we first review related work, and then discuss the specific model we consider in this paper. We then describe our formalism using the Kalman filter, and present an empirical validation. Finally, we conclude and discuss future work.

Related Work

A number of researchers have presented probabilistic computational trust models for single dimensional contracts. Jøsang & Ismail (2002) describe the Beta Reputation System whereby the reputation of a supplier is compiled from the positive and negative reports of agents who have interacted with it, and this reputation is represented by a beta distribution. Likewise, Teacy *et al.* (2005) use the beta distribution to describe an agent's belief in the probability that a supplier will successfully fulfill its commitments. They present a formalism based on Bayesian statistics that allows an agent (i) to estimate this probability from its own direct experience, (ii) to communicate these estimates as reputation reports using the sufficient statistics of the beta distribution, and (iii) to combine such reports to provide more accurate estimates.

While these models only deal with single dimensional contracts, other researchers have noted the need for multi-dimensional models, and indeed, a number of such models have been published. For example, both Sabater & Sierra (2001) and Griffiths (2005) present multi-dimensional trust models, in which agents form contracts based on multiple variables. Both models provide heuristics to update these dimensions given observations of contract outcomes, and to combine these dimensions into a scalar metric that can be used to select between suppliers.

In earlier work we have combined these approaches within a probabilistic multi-dimensional trust model in which the Dirichlet distribution is used to represent an agent's correlated beliefs regarding the probability that a supplier will successfully fulfill each contract dimension. We considered an agent that is attempting to estimate the expected utility of a contract, and showed that this leads to a principled means of combining multi-dimensional beliefs into expected utility.

In this paper, we significantly extend this approach by considering the case that the agents' direct experience represents contracts over heterogeneous dimensions. In this case, the formalism described above can not be used since it is not possible to simply aggregate the observed contract outcomes in this way. Thus, we must develop an alternative approach and here we use the Kalman filter in order to fuse these heterogeneous contract observations.

Expected Utility of a Contract

We start by considering an agent attempting to procure a bundle of services (such as audio, video and data services) from a single supplier. In order to make a rational decision, or to negotiate a price for this bundle, the agent must estimate the expected utility of a contract with this supplier. Thus, we denote the outcome of a contract as a vector, X, that indicates whether or not each service within the bundle was successfully delivered (e.g. $X = \{o_a = 1, o_b = 0, o_c = 0, \ldots\}$ indicates that service a was successfully delivered, while services b and c were not). If $u(o_a = 1)$ is the marginal utility that the agent derives if service a is successfully delivered[1], then the expected utility of the agent will depend on the probability that this happens, $p(o_a = 1)$. However, neither the probabilities, nor the correlations between them, are not known to the agent, and thus, it must use observations of previous contract outcomes to determine a

[1]Our formalism can be applied to more complex utility functions that exhibit complementarities between services, however for clarity we present the simpler additive example in this paper.

distribution over their possible values. It can then determine an expectation of the expected utility of the contract:

$$E[E[U]] = \hat{p}(X)^T U(X) \tag{1}$$

and a variance, describing its uncertainty:

$$\text{Var}(E[U]) = U(X)^T P(X) U(X) \tag{2}$$

where:

$$U(X) = \begin{pmatrix} u(o_a = 1) \\ u(o_b = 1) \\ u(o_c = 1) \\ \vdots \end{pmatrix} \tag{3}$$

Thus, the agent's estimate of the expected utility is dependent on a trust estimate composed of two expressions: a vector estimate of the probability that each service is successfully delivered:

$$\hat{p}(X) = \begin{pmatrix} \hat{p}(o_a = 1) \\ \hat{p}(o_b = 1) \\ \hat{p}(o_c = 1) \\ \vdots \end{pmatrix} \tag{4}$$

and a covariance matrix that describes the uncertainty and correlations in these estimates:

$$P(X) = \begin{pmatrix} V_a & C_{ab} & C_{ac} & \dots \\ C_{ab} & V_b & C_{bc} & \dots \\ C_{ac} & C_{bc} & V_c & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix} \tag{5}$$

where the diagonal terms, V_a, V_b and V_c, represent the uncertainties in $p(o_a = 1)$, $p(o_b = 1)$ and $p(o_c = 1)$, and the off-diagonal terms C_{ab}, C_{ac} and C_{bc} represent the correlations between these probabilities.

Heterogeneous Contracts

The previous section showed that in order to estimate the utility of a contract an agent must calculate a trust estimate composed of the vector, $\hat{p}(X)$, and covariance matrix, $P(X)$. In earlier work we presented a formalism using the Dirichlet distribution that allows an agent to calculate both these expressions from its direct experience of previous contract outcomes (Reece *et al.* 2007). Within this formalism, an agent that has observed N contract outcomes in total simply records, for each pair of services (e.g. a and b), the number of times that both were delivered successfully, n_{11}^{ab}, the number of times both were delivered unsuccessfully, n_{00}^{ab}, and both combinations in which one was successfully delivered and the other unsuccessfully delivered, n_{01}^{ab} and n_{10}^{ab}. These counts over contract outcomes can be communicated as reputation reports, and these reputation reports can be combined by simply aggregating the counts.

However, this formalism is limited to the case that contract observations are homogeneous (i.e. all agents observe contracts over the same dimensions). This is the case since there is no way of aggregating the counts over contract outcomes of an agent who has observed two services, with those of an agent who only observed one service. Thus, in this section, we present two formalisms that address the more general case where contract observations are heterogeneous. We first describe a simple benchmark formalism using independent beta distributions, and then describe our full formalism that uses the Kalman filter.

Inflated Independent Beta Distributions

We can provide a reasonable benchmark formalism for dealing with heterogeneous contracts through a simple extension of a single dimensional trust model. That is, we do not explicitly represent the correlations between the services within the bundle, but rather, we use independent beta distributions to represent each individual service. Thus, if an agent has direct experience of N previous contract outcomes, in which service a was successfully delivered n_a times, then the trust estimate, $\hat{p}(X)$, can simply be calculated using the standard result from the beta distribution[2] that:

$$\hat{p}(o_a = 1) = \frac{n_a + 1}{N + 2} \tag{6}$$

Similarly, we can calculate the diagonal terms of the covariance matrix, $P(X)$, by again using the standard result from the beta distribution that:

$$V_a = \frac{(n_a + 1)(N - n_a + 1)}{(N + 2)^2 (N + 3)} \tag{7}$$

Finally, rather than explicitly calculating the off-diagonal elements of the covariance matrix, we can derive a conservative covariance matrix[3] by simply setting the off-diagonal elements to zero, and multiplying the diagonal variance terms by the number of dimensions in the state vector, X. Thus in the case of two services we have:

$$P(X) = \begin{pmatrix} 2V_a & 0 \\ 0 & 2V_b \end{pmatrix} \tag{8}$$

This process is known as covariance inflation, and reflects the fact that while we do not know the correlations between the services, we know that they may be correlated, and thus, we require a conservative covariance matrix that covers any possible correlation (Hanebeck, Briechle & Horn 2001; Reece & Roberts 2005).

This simple formalism is attractive; by not explicitly modelling the correlations between services it allows us to fuse heterogeneous contract observations by simply aggregating the counts (i.e. finding the total number of times service a was successfully delivered out of the total number of contract observations). However, as we shall show later, the lack of explicit correlation information causes it to perform poorly. Thus, in the next section we develop a more sophisticated approach using the Kalman filter to fuse heterogeneous estimates containing correlation information.

[2] See Teacy *et al.* (2005) or Reece *et al.* (2007) for example.

[3] A covariance matrix is *consistent* (or *conservative*) when it is not less than the actual distribution of the true trustworthiness, $p(X)$, around the estimate, $\hat{p}(X)$. Conservative covariance matrices ensure that we never assign greater credibility to a trust estimate than it deserves, and are thus important when risk averse decisions are made (Uhlmann 2002).

A Kalman Filter Trust Model

The Kalman filter is a natural choice for our formalism, since within the academic literature of data fusion it is commonly used to fuse observations over multiple correlated dimensions (Bar-Shalom, Li, & Kirubarajan 2001). This work generally assumes Gaussian distributions. However, the Kalman filter can also be used for non-Gaussian distributions (Maryak, Spall, & Heydon 2004), and we describe how it can be applied to Dirichlet distributions in order to fuse heterogeneous trust estimates from multiple agents.

Our Kalman filter trust model operates by fusing an agent's prior trust estimate (calculated from an agent's own direct experience of previous contract outcomes) with reputation reports that are received from other agents in order to give a posterior trust estimate. As described earlier, these trust estimates are represented by a vector, $\hat{p}(X)$, and a covariance matrix, $P(X)$, and the standard form of the Kalman filter provides two equations to update these:

$$\hat{p}_{posterior} = \hat{p}_{prior} + K(o - \hat{p}_{prior}) \tag{9}$$

$$P_{posterior} = P_{prior}[1 - K] \tag{10}$$

where K is the Kalman gain:

$$K = P_{prior}(P_{prior} + R)^{-1} \tag{11}$$

and o is an observation with covariance R, that together represent the reputation reports received from other agents (we discuss the details of these later).

Now, when we have heterogeneous contracts, one or more dimensions of either the prior estimate or the reputation reports may be missing. Within the Kalman filter framework we can simply represent these missing contract observations by setting the corresponding diagonal elements of the covariance matrix to infinity. By doing this we are effectively saying that the estimate for this contract part has no certainty.

Actually, performing these matrix operations involving infinity can be problematic. We can avoid this by using the information form of the Kalman filter whereby an estimate is represented by its precision, Y, which is the inverse of the correlation matrix (i.e. $Y = P(X)^{-1}$), and its information estimate, $\hat{y}$, which is the product of the precision and the state estimate (i.e. $\hat{y} = P(X)^{-1}\hat{p}(X)$).

In this case, the missing information can be represented by inserting zeros into the precision matrix, and as before, the Kalman filter allow us to combine reputation reports with prior beliefs to yield a posterior information estimate and precision matrix:

$$\hat{y}_{posterior} = \hat{y}_{prior} + \hat{y}_o \tag{12}$$

$$Y_{posterior} = Y_{prior} + Y_o \tag{13}$$

where $Y_o = R^{-1}$ and $\hat{y} = R^{-1}o$. The information form of the Kalman filter is particularly useful within multi-agent systems since reputation reports from multiple agents are simply added (in any order) to an agent's prior estimate. However, the two forms are exactly equivalent, and we can easily switch between the two.

Thus having presented the Kalman filter in the context of a computational trust model, we describe how an agent's prior estimate is calculated from its own direct experience, and how other agents can communicate reputation reports calculated from their own direct experience.

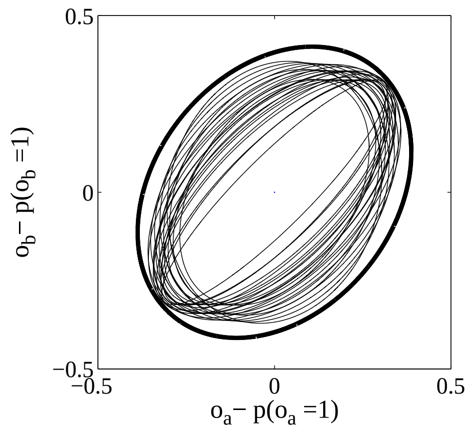

Figure 1: Conservative bounding matrix (*shown as a bold ellipse*), R^*, constructed from the family of all possible matrices, R (*shown as plain ellipses*).

Calculating a Prior Belief from Direct Experience: The prior belief of the agent is represented by a trust estimate, $\hat{p}(X)$, and a covariance matrix, $P(X)$. These can be calculated from an agent's direct experience using the Dirichlet formalism described in our earlier work (Reece *et al.* 2007). More specifically $\hat{p}(X)$ and the diagonal elements of $P(X)$ are calculated from the counts of contract outcomes (as per equations 6 and 7), while the full details of the Dirichlet distribution are required to calculate the off-diagonal terms of $P(X)$. See appendix A for full details.

This prior explicitly represents the correlations over the subset of services for which the agent has directly observed previous contract outcomes. When the agent has no direct experience of some services, it may simply insert infinity into the relevant diagonal element of $P(X)$ to reflect this lack of information (or alternatively insert zero into Y if the information form of the Kalman filter is being used).

Calculating Reputation Reports: The Kalman filter fuses a prior estimate with an observation, o, whose covariance is R. In our computational trust model, o and R together represent a reputation report and are calculated from the direct experience of the originating agent. This calculation is different from that which generates $\hat{p}(X)$ and $P(X)$, since the covariance R describes the variability of o about the true probabilities, $p(X)$, while the covariance $P(X)$ describes the variability of $p(X)$ about the estimate $\hat{p}(X)$. This is a subtle but important difference.

Calculating o is straightforward since it is a vector estimate of the probability that each service is successfully delivered (i.e. $o = \{o_a, o_b, o_c, \ldots\}$). It is calculated from an agent's previous contract outcomes, and thus, if the agent has observed N contracts in total, and service a was successfully delivered in n_a of these then $o_a = n_a/N$. Note that due to the reasons described above this expression is different from that shown in equation 6.

Calculating R is more complex. Since we are using the Kalman filter with a Dirichlet distribution (rather than the more common Gaussian distribution), the covariance, R, is itself dependent upon the probabilities that each service is successfully delivered, $p(X)$. These probabilities are not known; indeed, these are what we attempting to estimate. However, the beauty of the Kalman filter lies in its flexibility and we need not worry about finding R exactly. Provided that we can find a conservative matrix, R^*, to use in place of R, we can guarantee that our estimates will remain consistent. We can build such a conservative covariance matrix for R from an agent's direct experience and a commonly used method from the data fusion literature: namely covariance inflation (Hanebeck, Briechle & Horn 2001; Reece & Roberts 2005).

The full details of this calculation are presented in appendix B. However, we provide a sketch of the procedure here. Our starting point is the Dirichlet distribution over possible values of $p(X)$ calculated using the agent's direct experiences (as used above to calculate an agent's prior estimate). We then sample possible values of $p(X)$ from this distribution, and generate a family of possible covariance matrices for R. We then use covariance inflation to construct a conservative covariance matrix from the entire family. In figure 1 we present an example of this process for two services[4]. We plot the family of possible R matrices as ellipses, with the bounding conservative covariance matrix, R^*, in bold.

Example: To illustrate the formalism we consider an example where agent A is estimating the utility of procuring a bundle of two services, a and b. These two services share a common resource, and thus, the probabilities of them being successfully delivered are positively correlated. Agent A has some direct experience of procuring both services from this supplier, and thus, it can use the Dirichlet distribution to calculate prior estimates of $\hat{p}(X)$ and $P(X)$. This prior estimate is plotted in figure 2 as a dotted ellipse.

Agent B has also interacted with this supplier in the past, but in these interactions it has only observed contract outcomes involving service a. It communicates these contract outcomes to agent A in the form of an observation vector, o, and a conservative estimate of its covariance, R. This is shown as the dashed-dot ellipse in figure 2. Note that the variance in the b dimension is infinite (the covariance ellipse looks rectangular) reflecting the fact that agent B supplies no information about the reliability of service b.

Agent A can then fuse its own prior estimate with the reputation report received from agent B (using the information form of the Kalman filter and inserting the necessary zeros to indicate that agent B provides no information about service b). The resulting posterior estimate is shown as a solid ellipse in figure 2. Note that although agent B supplies information about service a only, the uncertainty in agent A's estimate for both dimensions is reduced. This occurs because agent A's prior estimate encodes a non-zero correlation between the services and the Kalman filter uses this to map the new evidence about service a onto service b.

[4]We choose two services for this and others examples in this paper since this allows us to plot covariance matrices as ellipses. Our formalism obviously generalises to any number of services.

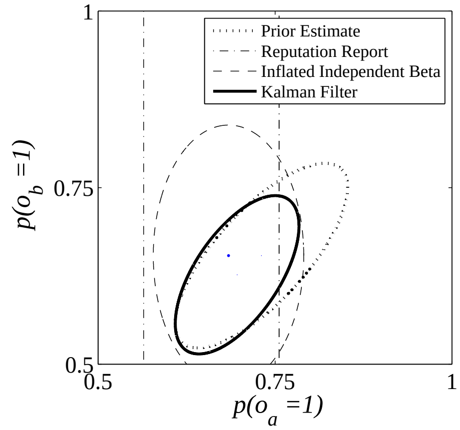

Figure 2: Kalman filter and inflated independent beta estimates for heterogeneous contract observations.

This is a key benefit of our formalism. We can compare this result to the benchmark presented earlier that uses independent beta distributions to describe each separate service (shown as the dashed ellipse in figure 2). Since this benchmark fails to represent the positive correlations between the services (i.e. the ellipse is not tilted to the right), it yields a highly pessimistic covariance estimate (i.e. the ellipse is substantially larger than that calculated by our formalism based upon the Kalman filter). In the next section we describe metrics that describe the information content and consistency of these estimates, in order to perform a more detailed comparison of these two formalisms.

Empirical Evaluation

In order to evaluate the effectiveness of our formalism, we present simulation results in which ten agents, each with their own direct experience of a supplier that provides two services, participate within a reputation system. We assume that one of these agents is attempting to evaluate the trustworthiness of the supplier in order to calculate the expected utility of interacting with it. As such, the agent must fuse its own direct experience with reputation reports received from the other nine agents. We compare two formalisms:

- **Inflated Independent Beta Distributions:** We use inflated independent beta distributions to represent each service separately (as described earlier).
- **Kalman Filter:** We use the formalism based upon the Kalman filter developed in this paper which explicitly captures correlations between the services.

In each simulation run, contract outcomes are drawn from an arbitrary joint distribution that induces correlations between the services. The contract outcomes are randomly allocated such that some agents observe both services, while others

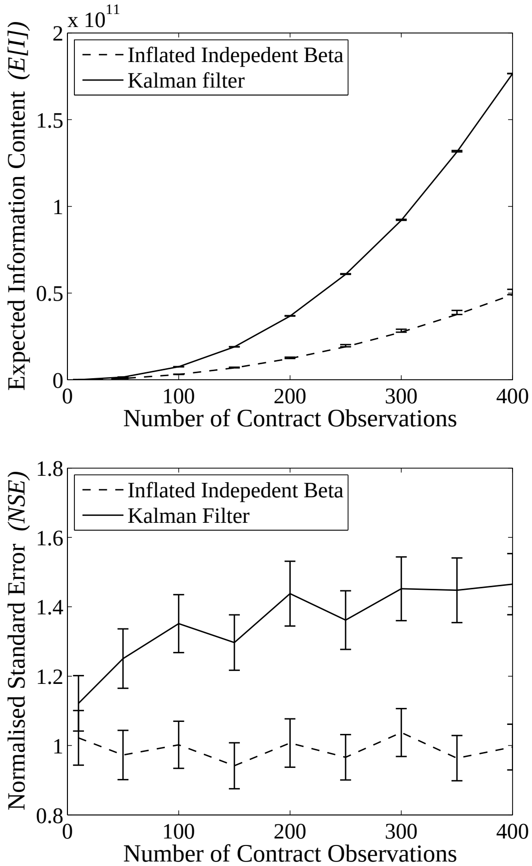

Figure 3: Comparison of the expected information content, $E[I]$, and normalised standard error, NSE, for formalisms using the Kalman filter and independent beta distributions.

observe just one service. We apply our formalisms to calculate posterior trust estimates and then calculate two metrics. The first is a scalar measure of the information content of the trust estimate; a standard way of measuring the uncertainty encoded within the covariance matrix (Bar-Shalom, Li, & Kirubarajan 2001). More specifically, we calculate the determinant of the inverse of the covariance matrix:

$$I = \det(P(X)^{-1}) \tag{14}$$

and note that the greater the information content, the more precise $\hat{p}(X)$ will be. The second metric measures the normalised error of the estimate:

$$E = [\hat{p}(X) - p(X)]^T \, P(X)^{-1} \, [\hat{p}(X) - p(X)] \tag{15}$$

We perform 1000 repeated simulation runs and calculate the expectation of these two metrics (and the standard error in these expectations). We note that the expectation of the normalised error is commonly termed the normalised standard error, NSE, and it describes the consistency of the estimate. A consistent estimate has a normalised standard error less than the cardinality of the trust estimate; two in this case. A normalised standard error much less than this value indicates that the covariance matrix is too conservative.

Method	$E[E[U]] \pm \sqrt{\text{Var}(E[U])}$
True Distribution	5.80 ± 0.27
Inflated Independent Beta	5.86 ± 0.53
Kalman Filter	5.82 ± 0.34

Table 1: Estimated expected utility and its standard deviation calculated from an agents posterior trust estimate.

In figure 3 we present these results (with the standard error in the expected values shown as error bars) as the number of contract observations ranges from 10 to 400. We note that the information content of the trust estimates generated by our Kalman filter formalism far exceeds that of those generated using inflated independent beta distributions (typically by a factor of three). By explicitly representing the correlations between the services our formalism generates more precise trust estimates. This increased precision is not realised at the cost of producing inconsistent estimates; the normalised standard error of both formalisms is less than two, and thus, they both generate consistent estimates. Finally, we note that as the number of contracts increases, the Kalman filter encodes more precise correlation information, and the difference between the formalisms also increases.

Finally, in table 1 we illustrate the effect that the precision of the trust estimate has on an agent's estimate of the expected utility of a contract (calculated using the relationships shown in equations 1 and 2 in an example setting where $u(o_a = 1) = 2$ and $u(o_b = 1) = 6$). While both formalisms generate estimates of expected utility close to the true distribution, the more precise covariance matrix of the Kalman filter results in a better estimate of the standard deviation of the expected utility (while that of the inflated independent beta distribution is approximately double the true value).

Conclusions

In this paper we addressed the need for a principled probabilistic model of computational trust that allows heterogeneous contract observations to be fused together. Prior to our work no such model existed. We considered the case of an agent procuring a bundle of services subject to correlated failures, and we showed that we could use the Kalman filter to fuse observations from agents who have direct experience of previous contracts for different subsets of these services.

Our future work concerns two areas. First, we note that the normalised standard error (*NSE*) of our formalism is generally much less than 2 (see figure 3). This suggests that there is some scope for deriving less conservative covariance matrices to represent the agents' reputation reports, and we are currently exploring this possibility. Second, we note that data fusion is a mature research field with many well developed techniques for filtering and fusing observations made by different agents. As such, we intend to incorporate some of these results within our computational trust model, and we are particularly interested in techniques to deal with erroneous or inconsistent estimates that are received from malicious (or misinformed) agents.

Appendices

A. The Dirichlet Distribution

In this section, we describe how an agent may use the standard results of the Dirichlet distribution to calculate the off-diagonal terms within $P(X)$. For each pair of service (e.g. a and b), we must consider all possible combinations of contract outcomes, and thus we define n_{ij}^{ab} as the number of contract outcomes for which both $o_a = i$ and $o_b = j$. For example, n_{10}^{ab} represents the number of contracts for which $o_a = 1$ and $o_b = 0$ (i.e. service a was successfully delivered, while service b was not).

Now, using the standard Dirichlet notation, we can define $\alpha_{ij}^{ab} \triangleq n_{ij}^{ab} + 1$ for all i and j taking values 0 and 1, and then, to calculate the cross-correlations between the two services a and b, we note that the Dirichlet distribution over pair-wise joint probabilities is:

$$\text{Prob}(p_{ab}) = K_{ab} \prod_{i \in \{0,1\}} \prod_{j \in \{0,1\}} p(o_a = i, o_b = j)^{\alpha_{ij}^{ab} - 1} \quad (16)$$

where:

$$\sum_{i \in \{0,1\}} \sum_{j \in \{0,1\}} p(o_a = i, o_b = j) = 1 \quad (17)$$

and K_{ab} is a normalising constant (Evans 1993). From this we can derive pair-wise probability estimates and variances:

$$E[p(o_a = i, o_b = j)] = \frac{\alpha_{ij}^{ab}}{\alpha_0} \quad (18)$$

$$V[p(o_a = i, o_b = j)] = \frac{\alpha_{ij}^{ab}(\alpha_0 - \alpha_{ij}^{ab})}{\alpha_0^2(1 + \alpha_0)} \quad (19)$$

where:

$$\alpha_0 = \sum_{i \in \{0,1\}} \sum_{j \in \{0,1\}} \alpha_{ij}^{ab} \quad (20)$$

and in fact, $\alpha_0 = N + 2$, where N is the total number of contracts observed. Likewise, we can express the covariance in these pair-wise probabilities in similar terms:

$$\text{C}[p(o_a = i, o_b = j), p(o_a = m, o_b = n)] = \frac{-\alpha_{ij}^{ab}\alpha_{mn}^{ab}}{\alpha_0^2(1 + \alpha_0)} \quad (21)$$

Finally, we can use the expression:

$$p(o_a = 1) = \sum_{j \in \{0,1\}} p(o_a = 1, o_b = j) \quad (22)$$

to determine the covariance C_{ab}. To do so, we first simplify the notation by defining $V_{ij}^{ab} \triangleq V[p(o_a = i, o_b = j)]$ and $C_{ijmn}^{ab} \triangleq C[p(o_a = i, o_b = j), p(o_a = m, o_b = n)]$. The covariance for the probability of positive contract outcomes is then the covariance between $\sum_{j \in \{0,1\}} p(o_a = 1, o_b = j)$ and $\sum_{i \in \{0,1\}} p(o_a = i, o_b = 1)$, and thus:

$$C_{ab} = C_{1001}^{ab} + C_{1101}^{ab} + C_{1011}^{ab} + V_{11}^{ab} \quad (23)$$

Thus, given a set of contract outcomes that represent previous interactions with a supplier, an agent may use the Dirichlet distribution to calculate an estimate of the probability that any service will be successfully delivered, $\hat{p}(X)$, and the uncertainty and correlations between these probabilities, $P(X)$, may be calculated and used as a prior belief in our Kalman filter trust model.

B. Covariance Inflation

In this section we describe a method whereby an agent can calculate a conservative matrix, R^*, to use in place of R, in the reputation report that it sends to other agents (see equation 11). The approach uses covariance inflation (Hanebeck, Briechle & Horn 2001; Reece & Roberts 2005) to calculate a bounding (i.e. consistent) covariance matrix from a (possibly sparse) set of contract outcomes. Covariance inflation is traditionally used to obtain a consistent covariance matrix for a family of covariance matrices when only the cross-terms in the family differ (i.e. the diagonal variance terms are assumed to be known). In this section we extend covariance inflation to the case where not only the off-diagonal terms, but also the diagonal variance terms, are unknown (but bounded).

We start by considering that the contract outcome observation covariance matrix, R, is given by:

$$R = \begin{pmatrix} p_a(1-p_a) & p_{ab} - p_a p_b & p_{ac} - p_a p_c & \dots \\ p_{ab} - p_a p_b & p_b(1-p_b) & p_{bc} - p_b p_c & \dots \\ p_{ac} - p_a p_c & p_{bc} - p_b p_c & p_c(1-p_c) & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix} \quad (24)$$

where p_κ, with $\kappa \in \{a, b, c, \dots\}$, is the probability that the service κ is successfully delivered (e.g. $p_a = p(o_a = 1)$ and $p_b = p(o_b = 1)$) and $p_{\kappa\nu}$ is the probability that services κ and ν, with $\kappa, \nu \in \{a, b, c, \dots\}$, are both successfully delivered (e.g. $p_{ab} = p(o_a = 1, o_b = 1)$).

The probabilities within this matrix (e.g. p_a and p_{ab}) are not known. However, they can be estimated by sampling the distributions (by observing contract outcomes) and then using the Clopper-Pearson method (Clopper 1934) to find *confidence intervals* over them. These intervals describe a family of possible covariance matrices R from which a conservative covariance matrix R^* can be calculated.

The Clopper-Pearson method calculates confidence intervals, I_κ and $I_{\kappa\nu}$, for each of the marginal distributions over p_κ and the joint distributions over $p_{\kappa\nu}$ respectively. The magnitude of the confidence interval is determined by a user specified *confidence value*. The confidence value is the probability that the interval contains the actual positive contract outcome probability. In many applications it is common to choose the 95 percentile confidence value. For any dimensions labelled κ and ν the confidence intervals are defined in terms of their upper and lower limits thus:

$$I_\kappa \triangleq [I_{\kappa,l}, I_{\kappa,u}] \quad (25)$$

$$I_{\kappa\nu} \triangleq [I_{\kappa\nu,l}, I_{\kappa\nu,u}] \quad (26)$$

Our aim is to find a covariance matrix R^* which is consistent with R for all values of p_κ and $p_{\kappa\nu}$ in their confidence intervals. For any covariance matrix R' to be consistent with R, we require that the diagonal elements of R' to be the largest possible values that the diagonal elements of R can take. Thus, we can restate our problem as that of finding a covariance matrix R^* which is consistent with all R' where:

$$R' = \begin{pmatrix} R'_{aa} & p_{ab} - p_a p_b & p_{ac} - p_a p_c & \dots \\ p_{ab} - p_a p_b & R'_{bb} & p_{bc} - p_b p_c & \dots \\ p_{ac} - p_a p_c & p_{bc} - p_b p_c & R'_{cc} & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix} \quad (27)$$

and, for all $\kappa, \nu \in \{a, b, c, \ldots\}$, $p_\kappa \in I_\kappa$ and $p_{\kappa\nu} \in I_{\kappa\nu}$ and $R'_{\kappa\kappa} = \max_{p_\kappa \in I_\kappa}\{p_\kappa(1-p_\kappa)\}$.

To determine the range of values that the cross-terms $c_{\kappa\nu} \triangleq p_{\kappa\nu} - p_\kappa p_\nu$ can take we use equations 25 and 26 to give the limits:

$$I_{\kappa\nu,u} - I_{\kappa,l} I_{\nu,l} \geq p_{\kappa\nu} - p_\kappa p_\nu \geq I_{\kappa\nu,l} - I_{\kappa,u} I_{\nu,u} \quad (28)$$

However, we can often find tighter bounds. Since R is a covariance matrix then the magnitude of the cross term $p_{\kappa\nu} - p_\kappa p_\nu$ must be less than the square root of the product of the variances $p_\kappa(1-p_\kappa)$ and $p_\nu(1-p_\nu)$. Since the variances are bounded above by $R'_{\kappa\kappa}$ and $R'_{\nu\nu}$ then:

$$|c_{\kappa\nu}| \leq \sqrt{p_\kappa(1-p_\kappa)p_\nu(1-p_\nu)} \leq \sqrt{R'_{\kappa\kappa}R'_{\nu\nu}} \quad (29)$$

Thus, we can obtain the range of possible values for the off-diagonal covariance terms as a function of the Clopper-Pearson intervals. The upper $U_{\kappa\nu}$ and lower $L_{\kappa\nu}$ bounds for each cross-term, $U_{\kappa\nu} \geq c_{\kappa\nu} \geq L_{\kappa\nu}$, are given by:

$$U_{\kappa\nu} = \min\{\sqrt{R'_{\kappa\kappa}R'_{\nu\nu}}, I_{\kappa\nu,u} - I_{\kappa,l}I_{\nu,l}\} \quad (30)$$

$$L_{\kappa\nu} = \max\{-\sqrt{R'_{\kappa\kappa}R'_{\nu\nu}}, I_{\kappa\nu,l} - I_{\kappa,u}I_{\nu,u}\} \quad (31)$$

To build the observation covariance matrix R^*, we iterate through the contract dimensions d. Initially, $R^*[1] = R'_{aa}$ where $R'_{aa} = \max_{p_a \in I_a}\{p_a(1-p_a)\}$. Each subsequent iteration $d = \{2, 3, \ldots\}$ augments a row and column to $R^*[d-1]$. Covariance inflation is used at each iteration to incorporate the unknown, but bounded cross-terms:

$$R^*[d] = \begin{pmatrix} (1+K_d S_d)R^*[d-1] & E_d \\ E_d^T & \left(1+\frac{S_d}{K_d}\right) R'_{dd} \end{pmatrix} \quad (32)$$

where $R'_{dd} = \max_{p_d \in I_d}\{p_d(1-p_d)\}$. The scalar S_d is calculated with the aid of two vectors, $C_d(l)$ and $C_d(u)$, that are of size $d-1$ and are obtained from the Clopper-Pearson confidence interval limits:

$$C_d(l) = [L_{1d}, L_{2d}, \ldots, L_{d-1\ d}]^T \quad (33)$$

$$C_d(u) = [U_{1d}, U_{2d}, \ldots, U_{d-1\ d}]^T \quad (34)$$

and is given by:

$$S_d = \max_v \left[\Theta \max\{\text{abs}(C_d(u) - E_d), \text{abs}(C_d(l) - E_d)\} {R'_{dd}}^{-1}\right] \quad (35)$$

where Θ is the sphering matrix of $R^*[d-1]$, $\max_v$ is the value of the maximum element in the vector, and $\max$ and abs are the element wise maximum and absolute operators respectively. The $d-1$ vector E_d and the scalar K_d are chosen to minimise the determinant (det) of R^*_d subject to the constraint that R^*_d is positive semi-definite. This constraint is a requirement of all covariance matrices. Thus the optimal value for $\{K_d, E_d\}$ is found using sequential quadratic programming (Fletcher 1987) to solve the nonlinear programming problem:

$$\min_{K_d, E_d} \det(R^*_d) \text{ subject to eigenvalues}(R^*_d) \geq 0 \quad (36)$$

The final result of this iterative procedure is a conservative covariance matrix, R^*, to use in place of R, in the agent's reputation report.

Acknowledgments

This research was undertaken as part of the ALADDIN (Autonomous Learning Agents for Decentralised Data and Information Networks) project and is jointly funded by a BAE Systems and EPSRC strategic partnership (EP/C548051/1).

References

Bar-Shalom, Y., Li, X.-R., and Kirubarajan, T. 2001. *Estimation with Applications to Tracking and Navigation*. Wiley Interscience.

Clopper, C., and Pearson, S. 1934. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 26:404–413.

Evans, M., Hastings, N., and Peacock, B. 1993. *Statistical Distributions*, John Wiley and Sons, Inc.

Fletcher, R. 1987. *Practical Methods of Optimization*, John Wiley and Sons, Inc.

Griffiths, N. 2005. Task delegation using experience-based multi-dimensional trust. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, 489–496.

Hanebeck, U. D., Briechle, K., and Horn, J. 2001. A tight bound for the joint covariance of two random vectors with unknown but constrained cross-correlation. In *Proceedings of the IEEE Conference on Multisensor and Integration for Intelligent Systems*, 85–90.

Maryak, J. L., Spall, J. C., and Heydon, B. D. 2004. Use of the Kalman filter for inference in state-space models with unknown noise distributions. *IEEE Transactions on Automatic Control* 49(1):87–90.

Ramchurn, S. D., Hunyh, D., and Jennings, N. R. 2004. Trust in multi-agent systems. *Knowledge Engineering Review* 19(1):1–25.

Reece, S., and Roberts, S. 2005. Robust, low-bandwidth, multi-vehicle mapping. In *Proceedings of the Eighth International Conference on Information Fusion*.

Reece, S., Rogers, A., Roberts, S., and Jennings, N. R. 2007. Rumours and reputation: Evaluating multi-dimensional trust within a decentralised reputation system. In *Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multiagent Systems*. In press.

Sabater, J., and Sierra, C. 2001. REGRET: A reputation model for gregarious societies. In *Proc. of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies*, 61–69.

Jøsang, A. J., and Ismail, R. 2002. The beta reputation system. In *Proceedings of the 15th Bled Electronic Commerce Conference*

Teacy, W. T. L., Patel, J., Jennings, N. R., and Luck, M. 2006. TRAVOS: Trust and reputation in the context of inaccurate information sources. *Autonomous Agents and Multi-Agent Systems* 12(2):183–198.

Uhlmann, J. K. 2002. Covariance consistency methods for fault-tolerant distributed data fusion. *Information Fusion* 4(3):201–215.