Trustworthy Service Selection and Composition – Reducing the Entropy of Service-oriented Web

Wei Liu

School of Computer Science and Software Engineering, the University of Western Australia, Crawley, Australia
e-mail: wei@csse.uwa.edu.au

Abstract— Current e-service selection relies upon service registries to make recommendations. It is almost impossible to capture and make use of valuable personal service evaluations. E-service composition on the other hand employs intelligent planning for composition plan generation, which fails to scale in a large open environment such as the Web. In this paper, we consider trust evaluation and propagation as the fundamental driving force for service selection and composition, and proposed a framework that will facilitate the growth and evolution of a service-oriented network into a self-organised trustworthy virtual collaboration environment. Viewing a service-oriented network (Grid or Web) as an ecosystem, this framework is based on personalised service recommendation, social network analysis and emergent intelligence.

Index Terms—trust, e-service, emergent intelligence, social network analysis.

I. INTRODUCTION

The enormous opportunities offered by the Web and Grid technologies propelled the desire for a visionary change of the document centric Web into an interoperable service backbone. Since Berners Lee et al. [1] unleashed this semantic Web vision in 2001, substantial research and development effort have been put into service-oriented infrastructure and supporting technologies. It also attracted spontaneous industry support because of the huge potential of leveraging existing virtual collaboration from data and information sharing to business process integration.

Current Web service technologies concentrate on interoperability, accessibility and publicity of atomic services. An atomic service is a standalone network accessible process that is independent of any other services. However, the real power of service-oriented computing lies in composing and integrating atomic services to perform complex tasks that can not be achieved by any single atomic service alone.

Unlike traditional distributed objects from which they grew out of, atomic services in a service-oriented paradigm are not centrally designed or software engineered to be collaborative as coherent and cohesive systems. Instead, they are designed to operate independently and may be dynamically composed at runtime. There is no one single system that can be referred to. Moreover, the openness of the environment allows for services to join and leave at anytime and anywhere. Self-interest behaviours are also quite possible and may even be very common because services can be developed by different vendors using different business strategies. The quality of the services can vary dramatically as well, some services are accurate and robust, and others might be error prone and unreliable. Without sufficient trust, the risk involved in exposing business data to untrustworthy trading partners or relying upon external service providers for mission critical tasks, becomes the main obstacle in the adoption of service-oriented architectures as the main-stream business transaction model. In other words, trustworthy service selection and composition becomes the critical research issue.

Current service selection is often done by the service registry, which only captures common Quality of Service criteria but not personal service evaluations based on direct experiences. Service composition on the other hand largely relies on intelligent planning for generating alternative composition plans. However, intelligent planning suffers from a close world assumption where all possible operations and processes are known a priori. This assumption limited the potential application of intelligent planning in such an open and dynamic environment like the Web.

This paper proposes a trustworthy service selection and composition framework that can potentially turn a service oriented network into a self-organised trustworthy virtual collaboration environment. Section II investigates the current research on trust and service composition. Section III describes the conceptual framework. Section IV proposes the key technological components enabling such a framework. The paper concludes in Section V with an outlook of the rich potentials of such a framework.

II. RELATED WORK AND PROBLEM STATEMENTS

Trust as a social phenomenon, as a mental state, as the basis for establishing business alliances and partnership has been widely studied in social network analysis [2], cognitive psychology [3], and economics. More recently, trust is taken as a high level social abstraction of agencies [3, 4] and studied in the multi-agent research communities. The research into establishing trust can be classified into two categories, centralised and distributed. An example centralised approach – trust management in distributed computing, assumes trusted third parties [5] to authorize only trusted agents to perform potential risky actions. Trust in this case is established by consulting an authority. Distributed approaches value direct and indirect past experiences and assume no authoritative third party. For example, reputation mechanisms used in economics evaluate trust according to direct past behaviours [6]. The assumption is that a party who has built up a good reputation over a long period of
time normally will not risk it by purposely misbehaving. In such settings, agents police themselves without the need of a trusted third party. Term soft security is used to describe reputation based trust evaluation. When no trusted third parties and no past experience are available, the indirect experiences or opinions of other agents can be consulted. This mechanism is called the referral networks approach, often used for locating services through trusted referrals [7]. At the very basic level, when there is no other agents around, no third parties and no direct past experiences, two parties establish trust by starting interactions with little risk and gradually build the trust [2].

The third party authority approach may work well for well-established close environment, for example, within an enterprise boundary. Companies like Amazon.com, eOpinions.com and eBay.com provide recommendations on books and electronic products etc to help users with product selection. The company collates ratings from anonymous reviewers and uses for recommendations. It works because the company has access to centralised information on what to sell (all products are registered), and to whom it may sell. However, for a large open environment where no single party can possibly have access to global information, such approach becomes no longer applicable. In the same token, current Web service infrastructure support on service registries\(^1\) falls into the same centralised category. The registries offer yellow pages service where service providers register their services and service consumers “flip through” the registries for matching services. The service registries can implement similar centralised technologies (e.g. collaborative filtering [8]) for providing recommendations. Such an approach only works if all service providers do take the trouble and register with one or more registries to advertise their services. However, to promote and attract wider participation in service development and consumption, the expectation on service developers has to be lowered and the service discovery process has to be simplified. Just like there is no need to advertise a Web page, there should not be any assumptions on voluntary service registrations. Instead, search engine like support should be provided for dynamic collection and indexing of services. The proposed framework will assume a purely distributed setting supported by a pilot service search engine.

Without centralised third parties in the assumption, we can take a referral network approach similar to Yolum and Singh [9]. Each service provider or consumer has a dedicated agent to represent it in a service social network. The representative agent maintains a neighbour list of other representative agents with whom it has direct interaction. Trust in such a network takes a meaning of multi-folds, because the representative agent can assume three different roles, as provider, consumer, or referrer. Depending on the different roles both parties play respectively, trust has different meanings. For example, a consumer’s trust on a provider is the expected Quality of Service (QoS) delivered by the provider. A consumer’s trust on a referrer is the referrer’s objective opinion on certain service provider’s QoS. In addition, trust is inherently a personal opinion [3]. Agent \(a\) likes some cheap motel service might be due to some hidden factors that are not explicitly specified in the standard QoS matrix. For example, the linen is fresh and the breakfast is nice. So it is insufficient to represent trust just as simple one dimensional scalars like most recommender systems do, a multi-dimensional representation is necessary to store details that can later be useful for recommendation. For example, agent \(a\) can give a full recount of why it prefers that motel to an inquiry agent \(b\). Agent \(b\) can make a judgement by matching its own preferences with agent \(a\). Real world metaphors can be found in the online consumer comments/feedbacks/testimonies of certain product or service. Here we will develop an adaptive trust representation ontology for agent understandable service testimonies.

Once we are able to select atomic service through trustworthiness evaluation, next level of service consumption is to integrate atomic services into composite services to accomplish higher level tasks. A composite service life cycle should include automatic generation, instantiation and execution of a composition plan. A composite service is an instance of a composite plan, when the component services are bounded by a set of selected service providers. Each service provider performs the role required by the component service. One composition plan can be fulfilled by different sets of service providers.

To the best knowledge of the author, very little attention was given to the composite service life cycle although languages describing the flow control of component services in a composition plan are abundant. This research is to fill the gap by streamlining the process of automatic generation of composition plans, binding the composition plan with sets of trustworthy service providers, and evolving certain composition plan instances into composite services.

The instantiation and creation of new composite services will trigger virtual collaborations between businesses at the process level. Intuitively, such virtual collaborations in the long run at the macroscopic level has the potential of turning an environment populated with scattered and isolated services into a more organised environment. According to the second law of thermodynamics, all ecosystems have a tendency of increasing their entropies. Maxwell’s standpoint of entropy as ordering sees that all ecosystems tend to go into disorder if no effort was put in to counter such tendency. In this framework, entropy theories can be used to explain and guide the self-organising behaviours of a service-oriented environment.

\(^1\) Microsoft, IBM, SAP and NNT-COM offer UDDI Business Registry nodes.
Rothchild [11] shows economy has a similar emergent structure as an ecosystem. Similar to an ecosystem and an economy, a service-oriented network is populated with unpredictable number of uncoordinated individual atomic services interacting with each other in a similar fashion as organisms do. So we propose that a service-oriented network can be modelled as an ecosystem, as shown in Fig. 1. We call it Network Economy. The openness of such an environment encourages the increase of entropy of the entire system. Countering such tendency is a challenging task. In fact, the standard languages for service description, composition and the directory registry services are all part of the efforts in an attempt to reduce the entropy of the Web. Such efforts have close counterparts in human society, e.g. natural languages, business languages, and yellow pages service correspondingly, which all play crucial roles in the civilized world. However, they only provide the enabling fundamentals, but not mechanisms for supporting the emergence of groups, local communities or virtual alliances – a characteristic of decreasing entropy in an ecosystem. The proposed framework will help study ways of engineering organisational behaviour into the service-oriented Web and use entropy as a measurement to evaluate the effectiveness of such approaches.

In summary, at the atomic service level, the research focuses on personalised trust representation for trustworthy service selection, answering: 1) What information needs to be captured and how to make the representation adaptive? 2) How is selection done without third party authorities?

At the composite service level, the research focuses on the creation of composite services, answering: 1) How to create composition plans without no closed world assumption? 2) How does a composition plan evolve into a composite service?

At the network economy level, we study the emergent self-organising behaviour, answering: 1) How to engineer individual entities so desired organisational behaviours, e.g. consumer communities, business alliances and trusted third parties can emerge from lower-level interactions? 2) How to measure the system entropy to evaluate the level of organisation?

III. Conceptual Framework

Fig. 2 shows an overview of the framework, consisting of two main components, Trustworthy Service Composition and Trustworthy Service Selection. Trustworthy service composition builds upon trustworthy selection. Trustworthy service selection is based on personalised service recommendation and distributed trustworthiness propagation. Traditional recommendation systems using Collaborative Filtering [8] assume centralised network as well as simple scalar ratings. In order to preserve the personalised feature of trust evaluations, we first need a richer representation framework of trust. The representation framework has to be flexible and adaptive to suit the diversified needs of agents. It also has to be exchangeable so other agents can consult for referrals. Therefore, we term this adaptive trust ontology. Then traditional collaborative filtering algorithms need to be extended to handle such multi-dimensional “rating” information. To allow trust opinions propagate among agents, referral networks are used to connect various types of agent together.

Once the service composition plan is constructed, trustworthy service composition will resort to trustworthy service selection for the component service binding. Trustworthy service selection recommends the most trustworthy services based on the requester’s anticipated QoS. Traditional symbolic AI approaches for service composition fail to scale in large open environments. In nature, large open systems – e.g. ecosystems succeed generations after generations by self-organising into highly ordered sophisticated systems from extremely simple independent individuals. We propose using simple service crawlers to do service clustering first and then make selections on the possible composition plan step. A team-oriented structure is used to balance the autonomy and authority to achieve better network level emergent structure. Marker information (pheromone) is used to identify popular valid composition path so new composite services can be born into the environment. The system entropy is measured to validate the effectiveness of such a composition process.

IV. Key Technological Components

A. Adaptive Personalised Trust Representation

The mental ingredients of trust, according to Castelfranchi and Falcone [3] include competence belief, disposition belief, dependence belief, fulfilment belief, willingness belief, persistence belief and self-confidence belief. This set of pertinent beliefs can be classified according to two basic types of services (information services and action-based services). Liu and Williams [12] have identified that in information services, only competence belief and sincerity belief are necessary to information services. Willingness belief and persistence belief etc are more applicable to world altering events in action-based services. Common QoS attributes and personalised evaluations can be incorpo-
ranged into the relevant pertinent beliefs.

Recent development in ontology languages provide rich reasoning framework for knowledge representation, ranging from formal logic based languages like OWL\(^2\) (Web Ontology Language), to less formal conceptual graph based languages like RDF\(^3\) (Resource Description Framework) and Topic Maps. Topic Map\(^4\) is a relatively new ISO standard. With three types of basic constructs (topics, associations, and occurrences), topic maps are capable of representing various relationships between concepts. The richness in the expressive power is comparable to conceptual graphs in artificial intelligence. Unlike other ontology languages, topic maps are relatively simple to implement. With the help of an XML specification on Topic Maps (XML Topic Maps – XTM), topic maps have better tool support compared to other languages. Schema-based mappings allow the conversion of knowledge captured via topic maps into other representations\([13]\), even when they represent different levels of abstraction. Moreover, work on visualizing topic maps has produced a large array of tools for representing and navigating of topic maps. The visualization of topic maps takes advantages of the human ability to navigate in two or three dimensions, recognize patterns, track movements, and compare objects of different sizes and colors\([14]\).

We propose to use XTM for representing different dimensions of the trust cognitive constituents and their relationships. Topic Maps give us the extensibility and adaptability for explicitly represent, query and visualise such a complex structure. For any knowledge representation mechanism to cope with update and change automatically, we have to take into account the possibility of redundancy and inconsistency as well. Truth-maintenance and belief revision in a multi-agent environment will be used together with the build-in Topic Map merging mechanisms to resolve possible conflicts in trust representation.

### B. Trustworthy Service Selection using Personalised Collaborative Filtering

Collaborative filtering is a common technique used by online trading system for product selection based on scalar ratings. With a richer trust representation, agents are able to record why a service gets selected but not the other, which is much informative than just scalar ratings. Semantic similarity matching can then be carried out in finding matches between the active consumer agent and the consumer base. We propose a new concept of personalised collaborative filtering to indicate the new way of making recommendation based on semantically rich similarity measures. Bernstein et.al.\([15]\) reported five different types of approaches for comparing complex objects in an ontology. These include ontology distance, information-theoretic approaches, vector space approaches, Levenshtein distance (i.e. edit distance) and full-text retrieval method. Therefore, the recommender will provide various similarity measures to matching the specified selection criteria from service requester with the recommender agent’s internal evaluation of the potential service providers.

### C. Trust Network Structural Establishment and Analysis as Multi-thread Referral Networks

Network structure analysis was carried out in a small scale agent simulated referral networks by Yolum and Singh\([9]\). Referrals are taken as endorsement and also as a way of selecting services in the absence of direct experiences. Aggregated trustworthiness is evaluated over sociability and expertise, high sociability indicates good referrals and high expertise gives rise to good service providers. Representing as directed acyclic graphs, the in degrees associate with the level of expertise and the out degrees associate with the level of sociability. Referral network structure analysis gives insights to how community forms and how referral centre emerge from the interaction.

However, Yolum and Singh\([9]\) take the view that referral agents are also consumer agents. Despite the validity of such a view, the meaning of the referral link becomes confusing at times, because it can represent both direct experiences with a service provider and a known neighbour consumer. To be able to keep track of various trust relationship, we need to have a clear view of the types of roles an agent can perform. An agent in a service-oriented distributed computing environment can assume three different roles, i.e., Consumer (C), Recommender (R) and Provider (P). The three roles do not have to be exclusive. An agent can perform one or more roles at the same time providing that the agent is capable of keeping track of individual communication channels. On the other hand, we can restrict the role an agent can play in one single business transaction to just one. In other words, in a single business process, an agent can only play one role, either as a consumer, a recommender or a provider.

The trust networks are therefore should vary according to the roles between two agents: Peer Trust (C-C, R-R, P-P), Quality Trust (C-P, C-R), Loyalty Trust (P-C, R-C), Reputation Trust (R-P) and Broker Trust (P-R). Each type of relationships in the referral network is then represented separately, so various types of emergence can be analysed. Peer trust should provide insights on alliance emergence. Quality trust gives rise to preferred services or recommenders, therefore trusted third party authorities. Loyalty trust can introduce special reward system to increase the bound between consumer agents and service providers. Reputation trust can potentially be useful when analysing the publicity and competency of certain services. Broker trust can be used by the service providers to check on the sociability of the recommenders and predicate how well their services can be reached by the consumers. We propose the concept of multi-thread referral networks for analysing the evolution of the service-oriented communities.

### D. Engineering Desired Emergent Communities

Typically, a service composition plan can either be obtained semi-automatically\([16, 17]\) through user interface or automatically through AI planning\([18, 19]\), rule based rea-

---

\( ^2\)http://www.w3.org/2004/OWL/

\( ^3\)http://www.w3.org/RDF/

\( ^4\)http://www.topicmaps.org/
soning [20], or linear logic [21]. However, AI planning and logical approaches have shown fundamental limitations because of the close world assumption. Such centralised top-down approach fails to scale and cope with the dynamic and open nature of the services oriented environment. The declarative view taken by symbolic AI is also difficult to map into the imperative process steps taken by the service specifications. We believe a paradigm shift is necessary. This proposed framework will provide support on an ecology-based service composition. Ant colonies sort out larvae, eggs, cocoons and food without running any sorting algorithm. Especially, each ant is an extremely simple entity and follows very simple reactive rules for actions. We can imagine that a consumer agent can implement or use such “ant algorithms” (service ant crawlers) to classify services into clusters as well. A comparison between the ant sorting algorithm and a proposed service clustering sorting algorithm is shown in Fig. 3 below.

<table>
<thead>
<tr>
<th>Ant Sorting</th>
<th>Service Sorting/Clustering</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Wander around the nest.</td>
<td>Search neighbourhood hood or repository.</td>
</tr>
<tr>
<td>2 Sense the presence of object.</td>
<td>Sense service of enough similarity.</td>
</tr>
<tr>
<td>3 Maintain short memory of what it has seen,</td>
<td>Maintain short memory for collected items.</td>
</tr>
<tr>
<td>4 If not carrying anything, probability of picking up sensed object decreases if it has recently encountered similar objects in the environment.</td>
<td>If nothing is in its registry, probability of copying the sensed service address to its registry decreases if recently encountered similar services in the environment.</td>
</tr>
<tr>
<td>5 If carrying something, probability of dropping it increases if it has recently encountered similar items in the environment.</td>
<td>If there is service address is in the registry, probability of removing it from the registry and adding it to the cluster record increases if it has recently encountered similar services in the environment.</td>
</tr>
</tbody>
</table>

Fig. 3 Pseudo Algorithm for Service Clustering

If no atomic service providers found, the crawling ant will attempt to construct a composition plan by clustering the available services according to their effects and rank them by trustworthiness. Similarity measures will be used to calculate the semantic distance between the desired and service effects, simple reactive rules similar to ant sorting behaviours can be used to sort services into effect-specific groups. The irrelevant ones can be ignored. Once the clusters are generated, the service providers in the intersection of all relevant clusters provide more effects than others. The input of these services will be used for another round of ant crawling until all input are known to the consumer agent or time-out conditions are reached so human users are involved for more input.

A service-oriented network evolves because lower level agents interact to survive. The multi-thread referral network described in Subsection C is used to establish connections between agents. A service provider remains alive when there still is some need for its services. New services are born out of composition plan instances when enough service consumers are using them. Such cellula automata [22] paradigm we take here will automatically collect distributed garbage services that are no longer connected with other agents in the network and provide mechanisms for supporting composite service lifecycle. Studies on the emergent behaviour cellular automata have shown that neither total order nor chaos is the best state for a complex system. Team oriented architecture provide a good balance between order and chaos because each team agent exerts certain guidance and supervision to ensure desired emergent behaviour, whereas each team member is able to exercise its own deliberation by trying out different alternatives based on the current state of the environment. Therefore, in terms of implementation, we will use team-oriented programming (e.g. JACK Teams\(^5\)). The team entity represents the consumer agent, while the team members are the crawling ants for each type of desired output of a component service.

A set of agents has autocatalytic potential if in some regions of their joint state spaces, their interactions cause system entropy to decrease. During the process of selecting and using component services, the crawling ants will leave pheromones (i.e. special tags that uniquely identify the composite service). The pheromone gets stronger when more crawling ants go pass the same route, a composite service is born when pheromone reaches certain level. The set of service providers involved in the process thus have autocatalytic potential.

Organisation level can be measured according to the entropy or the symmetry of a system. The higher the entropy or the higher the symmetry is, the lower the organisation level. Under this framework, system entropy can be measured on the system behaviours and the number of virtual communities emerged with certain time period. The effectiveness of the ecosystem approach can be quantitatively measured by the system entropy according to the entropy equation defined in [23].

V. CONCLUSION

The paper reports some ongoing work the author has carried out on studying the interplay of three key components of the future Web: trust, service composition and emergent behaviours in a network of large populations of autonomous entities. The innovative ecosystem view of a service-oriented environment allows the research to be carried out at three abstraction levels, the atomic service, composite service and network economy level.

Trust may have been theoretically studied in a diverse range of disciplines, but rarely can one find practical yet computational realistic framework for trust establishment.

\(^5\) JACK Teams is an agent oriented teamwork language, developed by Agent-Oriented Software, Ltd. Evaluation version available from http://www.agent-software.com/
and evolution, especially in the context of service-oriented computing. Build on personalisable adaptive representation of trustworthiness evaluation, the computational framework proposed by this research will extend traditional scalar rating-based collaborative filtering to context sensitive multi-dimensional collaborative filtering, which we term Personalised Collaborative Filtering (PCF). PCF is then used by referrer agents in referral networks for establishing connections between agents. The referral networks in this research extend traditional ones with richer relationships (we propose the notion of multi-thread referral networks) and are used to simulate and analyse how virtual alliances and communities can emerge from a large set of independent atomic services.

Such a framework will add critical components to the existing research at the atomic service level, including adaptive trust ontology, personalised collaborative filtering and multi-thread referral networks. At the composite service level, benefiting from the study on emergent behaviours in ecosystems such as ant colony, we model the service consumers as ant-like agents (service crawlers). The service crawlers are able to discover services, classify them into clusters, and construct paths between the clusters for multiple composition plans.

At the network economy level, entropy reduction is characterised by the number of communities emerged and the number of distinctive role each community can play. An entropy equation can be used to measure the effectiveness of the trustworthy based service selection and composition.

Acknowledgement:
We would like to thank the support from the Australian Department of Education, Science and Technology, through IAP grant #CG040014 which enables us to collaborate with EU project SATINE, and the University of Western Australia for both university and faculty research grants.

VI. References