2008 IEEE International Conference on Services Computing

SOAR: SOcially Aware Routing for Request Matching in Enterprise Environments

Asheq Khan
State University of New York, Buffalo
akhan6@cse.buffalo.edu

Hani Jamjoom
IBM Watson Research Center
jamjoom@us.ibm.com

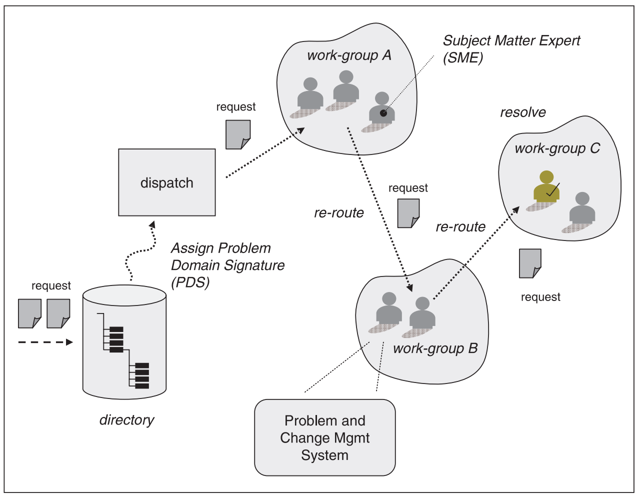

Figure 1. Overview of Service Management Environment

1 Introduction

Social networking has emerged as a powerful paradigm that connects people of similar interests. It has shown to be effective in routing queries in referral systems [3] and cellular networks [2], and also in the production of frequently asked questions (FAQ) repository [1]. In this paper, we extend its applicability to dispatch applications as part of fulfilling problem and change requests in outsourced IT environments.

As depicted in Figure 1, when a request is generated, it is assigned a *problem domain signature (PDS)* based on the symptom of the reported problem. A *problem domain* mirrors the various supported technologies (e.g., Lotus mail, Windows, etc). Using the signature, a dispatcher (or dispatching application) routes the request to a *work-group* that is best skilled to resolve the request. A work-group is basically a group of technicians—referred to as *Subject Matter Experts (SME's)*—who have been assigned to support a specific technology (e.g., email, Oracle, etc), typically for one or more outsourced accounts. That said, one of the primary challenges of request dispatch is to identify the correct work-group that can fulfill the request. It is the focus of our work.

Of course, it is not necessary that request routing is correct. In fact, with the exception of simple issues (such as password resets), requests often get re-routed and traverse multiple work-groups before being resolved as shown in Figure 1. However, the trail left behind resembles a thread of social interactions among SMEs in solving requests. We propose the SOcially Aware Routing (SOAR) system, which leverages these social interactions to improve the accuracy and efficiency of routing future requests. The system can incorporate multiple routing policies and provides load-balancing while maintaining a high degree of routing accuracy.

2 SOcially Aware Routing

The intuition behind SOAR is to represent requests originating from each PDS as a separate social network. Within a social network, each SME is assigned a weight based on his/her historical routing performances. As new requests arrive, SOAR chooses the appropriate SME to resolve the requests based on two routing policies (described next). The routing accuracy (in terms of SME performance) is continuously tracked in an online fashion to improve the accuracy of future routing decisions.

There are two aspects of SOAR that affects its efficiency: (1) *tracking metrics* and (2) *selection policy*. Tracking metrics capture the per-SME routing performance. We consider two metrics: (i) *inter-problem domain routing efficiency* and (ii) *intra-problem domain routing efficiency*. The former captures the connectivity/knowledge of an SME in routing requests across different problem domains. Such an SME, in theory, is able to contribute to multiple work-groups. Therefore, identifying well-connected SMEs can potentially increase the probability of locating the correct

978-0-7695-3283-7/08 $25.00 © 2008 IEEE
DOI 10.1109/SCC.2008.115

work-group for previously unseen PDS's. In contrast, intra-problem domain performance metric captures the effectiveness of an SME in correctly/quickly routing a request to a resolver within its problem domain.

We track the intra-problem domain routing efficiency in terms of the expected number of hand-offs/hops (E_{S_k}) for a request to be resolved once it passes through SME, S_k. For example, a knowledgeable SME will likely resolve a request himself and will have a very low E_{S_k}. On the other hand, an unskilled SME will have a much higher E_{S_k} as s/he is likely to forward a larger portion of requests to more knowledgeable SME's. In its simplest form, $E_{S_k} = \sum_{i=1}^{N} i \cdot P(i|S_k)$, where i is the number of hops from S_k to a resolving SME. P_i is the probability of taking i hop(s). Specifically, $P(i|S_k) = T_{hops|i,S_k}/T_{requests|S_k}$, where the numerator represents the total instances of requests taking i hop(s) from SME S_k to a resolver. The denominator corresponds to total number of requests handled by S_k.

Mirroring the above metrics, we consider two policies for routing requests to SMEs.

Policy A. The SME is selected based on his/her routing efficiency for a particular problem domain. The policy basically chooses the SME with minimum E_{S_k} for that problem domain. This policy can effectively reduce the request resolution time once a request is assigned to the correct work group as it favors skilled SMEs.

Policy B. The SME is selected based on his/her knowledge of multiple problem domains. That is, this policy chooses the SME with maximum social connectivity (from a social networking perspective) and can be considered as an information hub or gateway. This policy can be effective in correctly identifying an optimal SME for a new request whose assigned PDS does not map to any of the existing work-groups.

In order to balance the load across SMEs, SOAR tracks the number of requests routed to each SME. We consider load-balancing in terms of *(i)* a threshold on the frequency of requests routed to a SME and *(ii)* a round-robin distribution of requests among SMEs within the selected work-group(s). The former approach aims to prevent overburdening a SME with requests within any period of time. On the other hand, the latter scheme does not take into consideration the current load of a SME in routing requests. SOAR ensures that the load-distribution preserves a high degree of accuracy in identifying the correct resolver and/or minimizing the request resolution time.[1]

3 Evaluation

The proposed routing system is tested and validated using a large data set containing over 2 million problem and change records for a large number of outsourced accounts.[2] We replay the requests to mimic the evolution of the underlying system and consider three cases.

Case A. The PDS of a new request matches a single work-group and thus, only the most suitable SME (S_k) within the group needs to be selected. Under policy A, S_k has the minimum E_{S_k} value among all SMEs within the matched group. Under policy B, the chosen S_k has the highest connectivity/membership among all SMEs within the matched group.

Case B. The PDS of a new request matches multiple work-groups. This typically happens because PDS's are too coarse grain and may describe multiple technologies supported by different groups. In this case, a weighted probability for each matched group is first computed. These weights are computed from historical requests[3] and are the normalized occurrence frequency of work-groups for each PDS. Using these weights, the routing efficiency metrics for all SMEs in all matched groups are compared. Under policy A, the SME with the lowest metric is selected. On the other hand, the SME with the highest weighted connectivity is chosen as per policy B.

Case C. The PDS does not match with any of the work-groups in the routing table. In this case, under policy B, the SME with the highest connectivity across all the work-groups in the system is chosen as s/he is likely to act as a good relaying node.

For all test cases, it is ensured that the next hop SME is chosen with a certain degree of confidence. The degree of confidence is measured by the number of the requests an SME has previously handled. This filters out SMEs who have handled a fewer number of requests which, may result in a misleadingly low E_{S_k}.

Based on preliminary evaluation, our scheme is found to attain a dispatch accuracy of 91% in the steady state. We are in the process of performing more detailed evaluation and considering a limited pilot in a production environment.

References

[1] J. Davitz and et. al. ilink: Search and routing in social networks. In *ACM KDD*, 2007.

[2] H. Tomiyasu and et. al. Profile-based query routing in a mobile social network. In *IEEE Int'l Conference on Mobile Data Management (MDM)*, 2006.

[3] B. Yu and M. P. Singh. Searching social networks. In *Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2003.

[1] Details of dispatch load balancing are omitted for space considerations

[2] Due to the sensitivity of the data, the exact number of accounts cannot be disclosed.

[3] We use different information aging schemes to account for the changes in the underlying dataset.