



RESUME DE THESE DE DOCTORAT

Nom & Prénom(s)	Mostefai Abdelkader
E-mail (obligatoire)	Mostefaia_aek@yahoo.fr
Spécialité	Informatique
Titre	Rétro ingénierie des Systèmes d'Information pour une Architecture Orientée Services
Date de soutenance	25/06/2014
Nom, prénom(s) et grade de l'encadreur	Professeur : Malki Mimoun

Résumé :

L'une des caractéristiques fondamentales des logiciels est le besoin d'évoluer pour répondre aux nouvelles spécifications. La première loi de Lehman [M.M. Lehman and L. Belady, 1985] illustre ce fait : « Un programme utilisé dans un environnement du monde réel doit nécessairement changer, sinon il deviendra progressivement de moins en moins utile dans cet environnement ».

Actuellement, l'architecture orientée services (noté SOA) et la technologie des services web sont des approches prometteuses pour réaliser une entreprise agile et mettre en place une flexibilité durable. Mais la flexibilité et l'agilité de l'entreprise dépendent aussi de celles de son système d'information dont les composants constituent des applications patrimoniales héritées du passé informatique (ou applications légataires). En effet ces applications sont inflexibles, complexes et de mauvaise qualité. De plus, elles sont écrites avec des langages obsolètes et sont mal documentées. Mais elles représentent souvent des investissements considérables. L'entreprise préfère alors migrer ces applications (ou aussi logiciels) vers l'architecture SOA par un processus de réingénierie. La migration des systèmes information vers les architectures orientées services exige l'identification du code existant qui est capable de mettre en œuvre un nouveau service.

Dans notre travail de thèse, nous proposons trois approches complémentaires pour identifier un service à partir du code source des dits logiciels. La première est basée sur les algorithmes génétiques et la cohésion fonctionnelle pour extraire à partir du code source une implémentation quasi complète jouant le rôle d'un service donné. Dans la seconde approche, nous proposons une solution basée sur la récupération de l'architecture logicielle qui est extraite par un processus de classification non-supervisé en utilisant un nouveau modèle, nommé Sac à scores de similarité (BoSS). Ce modèle qui est basé sur une ontologie de domaine facilite énormément, à base de l'extraction de l'architecture fonctionnelle, la découverte efficace des services. La troisième approche est une approche basée sur les techniques de RI (Recherche Information), les ontologies, et l'analyse formelle de concepts (FCA) pour identifier efficacement des services d'un logiciel.

Des expérimentations ont été menées, afin d'évaluer la validité des approches proposées. Les résultats montrent que ces approches localisent efficacement un service(s) dans un logiciel légataire.

Mots clés : Web services ; Architecture Orienté service, rétro ingénierie, Logiciel légataire, Cohésion, Algorithme génétique, Architecture logiciel, analyse formelle de concepts, Recherche d'Information

Abstract

Legacy software continuously evolves to new requirement in business, technical and operating environments. Lehman's law of continuing change [M.M. Lehman and L. Belady, 1985] states this fact: 'Systems must be continually adapted or they become progressively less satisfactory'.

In nowadays, service oriented architecture (SOA) and its related technologies, web services, are a promising approaches to build an agile enterprise to solve challenges of globalisation and mobility.

RESUME DE THESE DE DOCTORAT

But Enterprise agility and flexibility depend on the flexibility and the agility of the enterprise information systems formed by its software applications (i.e., Legacy software). In fact, these applications are too complex, inflexible and have a degraded quality. Furthermore these applications are written with outdated languages and lack the proper documentations that describe correctly these applications. Despite the fact that these applications posse a bad properties, it seems that are of great interest for the enterprise. Enterprise prefers migrating of these applications towards SOA. Migrating legacy systems towards SOA requires the identification of legacy code that is able to implement the new services.

In this thesis, we propose three complementary approaches to identify services from legacy applications. The first approach use genetic algorithm (GA) and functional cohesion metric (FCOM) to search for the approximate implementation solution to this problem. The second one, use the architecture extracted from the legacy software to derive the services implementation. We propose the Boss model (Bag of Similarity Scores), a model to represent software entities. This model is ontology based and is personalised to guide the derivation of services from the legacy software. The third one uses information retrieval techniques, formal concept analysis and ontology to derive an approximate implementation of the services.

A set of experiments are done and show that these approaches identify services in an efficient manner.

Keywords : Reengineering, Reverse engineering, Legacy software, SOA, Web services, Ontology, Wrapping, Boss, Architecture, Information retrieval, Genetic algorithm, FCA.

ملخص

واحدة من الخصائص الأساسية لأي برمجيات هي خاصية التطور لمواكبة التغيرات التقنية و تغير المواصفات. ليان وبلادتي [ليمان و بلادتي، 1983] يصرح بذلك "البرمجيات يجب أن تتطور أو تصبح تدريجيا غير مرضية".

حاليا بنية الخدمية أو الخدمات (سرفيس أوريننتد أرشكتكثير) و تقنياتها، خدمات الواب، أصبحت واحدة من المقاربات الأكثر استعمالا لصناعة مؤسسات رشيقة لحل مشكلات التواصل بينها. لكن رشاقة المؤسسة و مرونتها مربوط بمدى مرونة نظامها المعلوماتي. هذا النظام المشكل من برمجيات المؤسسة و المسمى البرامج الموروثة. هذا النظام نظام بني بتقنيات قديمة مثل كوبول، الجافا. الخ. مما يجعله غير مرن و رشيق و يمنع بذلك تجسيد الرشاقة المرجوة داخل المؤسسة. ومع ذلك أيضا فهو يرى على انه كامل بالنسبة للمؤسسة لأنه ببساطة يحقق ما تريده بدون أخطاء. لذلك أحسن حل للمؤسسة هو تهجير هذه البرمجيات إلي بنية الخدمية باستعمال تقنية الريانجينيرنغ و الريفرس انجينيرينغ. لتحقيق هذا التهجير يجب أن نحدد أولا ماهي الخدمات الموجودة في هذه البرمجيات لإعادة استعمالها في بنية الخدمية.

هذه المذكرة تقدم ثلاث مقاربات متكاملة لتحديد الخدمة التي سيعاد استعمالها حسب قواعد بنية الخدمات. الأولى تستعمل الخوارزميات الجينية و الترابط الو ضيفي لإيجاد الخدمة. الثانية تستعمل البنية المستخرجة للبرنامج الموروث عن طريق عملية التكتل المبنية على أساس النمط المقترح "بوس: سلة قيم التشابه". النمط بوس مؤسس على أساس أنتولوجيا المجال الخاص بالبرنامج الموروث. الثالث يستعمل تقنيات: البحث المعلوماتي، الانتولوجي و تقنية تحليل الكونسابت لتحديد الكود الذي يجسد الخدمة.

كل هذه المقاربات قد جربت على برمجيات مختلفة و أثبتت فاعليتها.

كلمات مفتاحية: الريانجينيرنغ ، الريفرس انجينيرينغ، بنية الخدمات ، خدمة الواب، البرامج الموروثة، الانتولوجي، البحث المعلوماتي، الخوارزميات الجينية ، تحليل الكونسابت، البنية، النمط بوس.



FACULTE DE TECHNOLOGIE

RESUME DE THESE DE DOCTORAT