УДК 519.872, 519.688, 681.63

Исследование эффективности применения моделей доверия на основе репутации в Grid-системах

Куссуль О.М., Новиков А.Н., Швец С.С.
Национальный технический университет Украины «Киевский политехнический институт»
olgakussul@gmail.com

Abstract

Kussul O., Novikov O, Shvets S. Analysis of Efficiency of Reputation-based Trust Models Application in Grid-systems. *This paper describes main approaches to the development of reputation-based trust models for Grid systems. The extension of the existing utility-based reputation model for Virtual Organizations in Grid systems is proposed. Experiments were run to verify adequacy and efficiency of the proposed approach.*

Введение

С развитием электронной коммерции в Интернет доверию стали уделять повышенное внимание. Клиенты должны доверять продавцу, поскольку передают ему личные данные, а продавец должен доверять клиенту для того, чтобы предоставлять ему свои услуги. В Grid-системах ключевой идеей является совместное использование ресурсов [1], поэтому возникает необходимость во взаимном доверии пользователей и поставщиков ресурсов. В Grid-системах небольшого размера все участники находятся в отношении полного доверия. Например, в Украинском Академическом Grid-сегменте все участники принадлежат к НАН Украины, и на этом основании возникает полное доверие. Но в более масштабных Grid-системах участники зачастую могут быть напрямую не связанны друг с другом, и существует риск того, что кто-то из участников окажется недобросовестным и злоумышленным. Уменьшить эти риски и призваны механизмы доверия. На сегодняшний день уже существует множество концепций доверия и их реализаций [2, 3, 4, 5, 6].

В работе [7] показано, что доверие - это высокоэффективная технология, и ее внедрение позволит обезопасить электронные транзакции. При этом доверие описывается как важный и сложный предмет, связанный с честностью, правдивостью и надежностью доверенной особы или сервиса. Однако единой формулировки понятия доверия так и не достигнуто [8]. Приведем основные два определения доверия.

Когда мы говорим, что доверяем кому-то или что кто-то заслуживает доверия, то мы неявно имеем в виду, что возможность выполнения действия, которое будет полезным или как минимум не нанесет ущерб нам, является достаточно высокой, чтобы вступить с ним в определенные отношения [9].

В работе [10] доверие определяется как предел, до которого одна сторона хочет полагаться в определенной ситуации на что-то или кого-то с ощущение относительной безопасности, даже если возможны негативные последствия.

Актуальность управления доверием в Grid

В общем случае, целью механизмов безопасности является обеспечение защиты от злоумышленников. При этом различают два подхода [11]: „жесткая безопасность”, которая используется для описания традиционных механизмов, таких как аутентификация и контроль доступа; „мягкая безопасность” для, так называемых механизмов социального контроля, примером которых и является доверие. Доверие определяется в терминах отношений между доверителем (субъектом), который что-то доверяет, и доверенным лицом, которому что-то доверяют. Как отмечается в работе [10], традиционные механизмы безопасности („жесткая безопасность”) не могут предоставить защиту от угроз, вызванных злоумышленными ресурсами, или от злоумышленных действий авторизованных пользователей. В свою очередь, системы доверия могут предоставить защиту от таких угроз, которые являются специфичными для Grid-систем.

Ключевой концепцией Grid-сообщества являются виртуальные организации. Виртуальная организация (ВО) – это временное или постоянное объединение географически распределенных отдельных особ, групп, подразделений организаций или целых организаций, которые делятся ресурсами, возможностями и информацией для достижения общих целей [1]. Организация

OGF (Open Grid Forum) инициировала построение нового поколения программного обеспечения среднего уровня путем расширения современных технологий Web-сервисов в рамках Open Grid Services Architecture (OGSA) Рабочая группа OGF OGSA Security Workgroup выделила следующие актуальные задачи обеспечения безопасности в Grid:

– решения, направленные на интеграцию; в этом случае существующие сервисы и интерфейсы необходимо сделать более абстрактными для обеспечения расширяемой архитектуры;

– решения, направленные на интероперабельность с тем, чтобы сервисы, принадлежащие различным организациям с разными механизмами и политиками безопасности, могли взаимодействовать друг с другом;

– решения, направленные на определение, управление и внедрение политик доверия в динамические Grid-организации.

Виртуальные организации динамически формируются, существуют некоторое время и распадаются. Поэтому эффективность их работы зависит от доверия. В простом случае, когда одна сторона ручается за другую, вопрос доверия решается на основе «личного контакта». Другим примером таких «личных контактов» является вариант, где авторизированная организация выдает сертификаты. Однако такие «личные контакты» не масштабируемы в случае нетривиальных ВО. Для этого необходимы другие технологии, основанные на управление доверием на основании репутации с тем, чтобы создавать и проводить мониторинг таких ВО.

Существующие подходы управления доверием в Grid

В современных Grid-системах доверие реализуется путем использования механизмов безопасности [12]. Существующие механизмы обеспечивают одноразовую аутентификацию, которая основана на сертификатах, гарантирующих принадлежность узла к доверительной организации. Если какая-то организация желает присоединиться к некоторой ВО, она должна выполнить требования сертификационного центра (СЦ). Этот процесс обычно не обходится без участия человека. Поскольку Grid-системы постоянно развиваются и расширяются, возникает необходимость оценки и управления доверием организаций, участвующих в Grid-вычислениях. На сегодняшний день такие модели доверия строятся на основе оценки репутации сущности и существуют в основном для других прикладных областей.

В общем случае, под репутацией подразумевается мера надежности. Посредством репутации можно построить доверие от одной сущности к другой. Согласно [13], репутация – это предположение о поведении агента на основе имеющейся информации или наблюдений о его поведении в прошлом. В таком случае, для оценки репутации необходимо наличие данных о поведении агента в пошлом. Можно выделить следующие свойства систем обеспечение доверия на основе репутации:

– метрики доверия и репутации: обычно используют значения в диапазоне -1...+1, или 0...1. В последнем случае значение соответствующей метрики может интерпретироваться как вероятность;

– тип обратной связи: информация о репутации может быть позитивной или отрицательной. Некоторые системы основаны на сборе информации любого типа, в то время как другие системы только одного. Например, после выполнения транзакции данные о репутации могут быть представлены в бинарном, дискретном или непрерывном виде;

– надежность: модель доверия должна защитить пользователей от злоумышленной информации, в том числе неверных метрик доверия или репутации, распространяемой внутри системы другими пользователями.

Что касается непосредственно Grid-систем, следующие два свойства являются особо важными:

– обсуждение соглашения об уровне услуг (SLA) или качестве обслуживания (QoS): модели доверия должны учитывать выполнение данных соглашений и обеспечения необходимого качества;

– агрегирование доверия: это особенно важно в концепции ВО. Здесь актуальны две задачи: получение уровня доверия ВО в целом на основе уровня доверия ее участников и оценка уровня доверия некоторой организации, участвующей в определенной ВО.

Модели репутации могут обеспечить в Grid большую надежности путем решения проблемы устойчивости к отказам или путем улучшения распределения ресурсов и планирования. В любом случае, использование моделей на основе репутации дает возможность системе создавать «мягкую» версию доверия. На сегодняшний день существуют несколько подходов к построению моделей, основанных на репутации для Grid-систем.

PathTrust [14] – это система репутации, предложенная для выбора членов ВО на этапе формирования. Для того чтобы войти в процесс формирования ВО, организация должна зарегистрироваться с инфраструктурой сети предприятия (enterprise network) путем предоставления некоторых сертификатов. Помимо управления пользователями, сеть предприятия предоставляет централизированный сервис репутации. Когда разрушается виртуальная организация, каждый член оставляет определенные значения обратной связи для сервера репутации для других членов, с кем он вступал во взаимодействие. Эти значения обратной связи могут быть положительными или отрицательными. Система требует, чтобы после каждого

взаимодействия участниками были выставлены оценки. PathTrust является одной из первых попыток применить методы репутации к Grid-системам и подходы к управлению ВО. Однако при этом в этой системе не рассматриваются организационные аспекты. Предложенной модели не хватает динамики, так как обратную связь собирают только в момент разрушения ВО.

В работе [15] решается задача оценки уровня доверия между агентами мультиагентной системы, используя данные непосредственных наблюдений и информации о репутации. В отличие от существующих подходов, которые основаны на использование эвристик, предложенный подход HABIT (Hierarchical And Bayesian Inferred Trust Model) использует статистические байесовские методы. В частности, модель HABIT не ограничена к способам описания поведения агентов и в общем случае может быть адаптирована для предсказания как дискретного, так и непрерывного описания. Еще одним преимуществом данной модели является возможность получения оценок доверия, которые вычисляются путем поиска корреляций с поведением групп известных агентов. Это особенно важно в тех случаях, когда у доверителя нет предыдущего опыта или данных о репутации. Такая ситуация возникает для новых агентов, и во многих методах полагается, что уровень доверия для таких агентов минимален, что не всегда верно. Но данная модель трудно реализуемая в реальной распределенной системе, а также является еще недоработанной.

В работе [16] предложена модель доверия, основанная на модели репутации для ВО, которая может быть использована для оценки пользователей, ресурсов или поставщиков ресурсов. Эта модель основана на вычислении функции полезности, которая выражает удовлетворенность одного объекта его взаимодействием с другими объектами, принимая во внимание ключевые качества свойственные оцениваемому объекту. Для оценки репутации ресурсов эта модель доверия требует наличия системы мониторинга для сбора данных о качестве обслуживания. Для определения функции полезности вводится договоренность об уровне обеспечения качества услуг между пользователем и поставщиком для конкретного ресурса в ВО. Фактически, это ожидаемое качество обслуживания на ресурсе. Функция полезности определяется путем сопоставления реальных результатов мониторинга и ожидаемым уровнем обеспечения услуг. Таким образом, репутация каждого ресурса формируется в соответствии со значением функции полезности, которые агрегируются для одного ресурса и всех пользователей, которые с ним взаимодействовали. Аналогично, оценивается репутация пользователя: однако для определения функции полезности вместо договоренности об уровне предоставленных услуг используются заданные правила поведения пользователя на ресурсе, которые сопоставляются с результатами мониторинга реальной деятельности. Такая модель может быть адаптирована для Grid-систем, поскольку не требует прямого ответа от пользователя, а функция полезности является мерой удовлетворенности пользователя. Однако в данной работе не рассматривается вопрос, как устанавливается начальный уровень доверия для новых ресурсов или пользователей.

Модель репутации для виртуальных организаций на основе вычисления функции полезности

Модель репутации, предложенная в данной статье основанная на модели предложенной в [16]. Основные изменения относятся к модели репутации для пользователей ВО.

Основные определения. Центральным понятием для моделей репутации является организация [16]. Организация предоставляет ресурсы, а также существуют пользователи, которые принадлежат к данной организации (см. рис. 1).

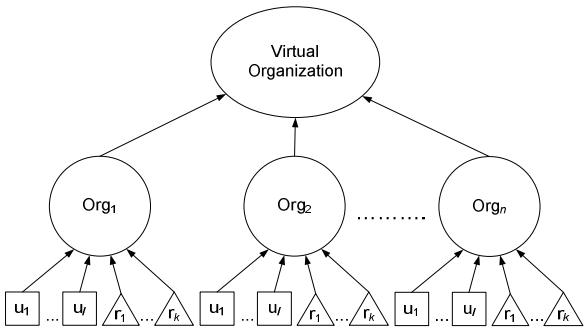

Рисунок 1 – Соответствие между ВО, организацией, пользователями и ресурсами

Вследствие этого, организация может быть описана следующим образом:

Определение 1. Организация представляется в виде следующего набора:

$$Org = \left\{ o\_id, \bigcup_i r_i, \bigcup_j u_j \right\}, \qquad (1)$$

где o_id является идентификатором организации, $\bigcup_i r_i$ и $\bigcup_j u_j$ являются соответственно ресурсами и пользователями принадлежащими к этой организации.

Все существующие организации обозначим $\bigcup_n Org_n$.

Виртуальную организацию (ВО) можно представить, как набор организаций. Организации объединяют свои ресурсы на временной или постоянной основе для достижения совместных

целей [1]. Следует отметить, что в общем случае организация может предоставлять ВО только подгруппу своих ресурсов, а также один ресурс может использоваться разными ВО. Тоже относится и к пользователям организации.

Определение 2. Представим виртуальную организацию в следующем виде:

$$VO = \left\{ vo\_id, \bigcup_k r_k, \bigcup_l u_l, f_{vo\_id}(), g_{vo\_id}() \right\}, \quad (2)$$

где *vo_id* — идентификатор ВО; $\bigcup_k r_k$ и $\bigcup_l u_l$ — соответственно ресурсы и пользователи различных организаций, которые принимают участие в функционирования ВО; $f_{vo\_id}()$ и $g_{vo\_id}()$ — функци принадлежности, определяемые следующим образом:

$$f_{vo\_id} : \bigcup_k r_k \rightarrow \bigcup_n Org_n \text{, т.к. } f_{vo\_id}(r_k) = o\_id, \quad (3)$$

$$g_{vo\_id} : \bigcup_l u_l \rightarrow \bigcup_n Org_n \text{, т.к. } g_{vo\_id}(u_l) = o\_id. \quad (4)$$

В общем случае эти функции могут быть реализованы с использованием сервисов сервера управления ВО (Virtual Organisation Management Server — VOMS). Используя эти функции, можно получать любую требуемую информацию о членстве организаций, ресурсов и пользователей в ВО. Например, набор ресурсов, предоставляемый организацией с идентификатором *o_id* в определенной ВО с идентификатором *vo_id* определяется:

$$\{ r \in \bigcup_k r_k : f_{vo\_id}(r) = o\_id \}. \quad (5)$$

Обозначим этот набор следующим образом: $f^{-1}{}_{vo\_id}(o\_id)$.

Аналогично можно перечислить всех пользователей из организации с идентификатором *o_id* участвующих в ВО с идентификатором *vo_id*:

$$\{ u \in \bigcup_l u_l : g_{vo\_id}(u) = o\_id \} \equiv g^{-1}{}_{vo\_id}(o\_id). \quad (6)$$

Предположим, что нам нужен список всех организаций, из $\bigcup_n Org_n$ которые предоставляют ресурсы определенной ВО с идентификатором vo_id, или чьи пользователи принимают участие в этой ВО. Согласно (3), (4):

$$\{ Org \in \bigcup_n Org_n : \text{если } \exists\ r \in \bigcup_k r_k \text{, то } f_{vo\_id}(r) = o\_id \}, \quad (7)$$

$$\{ Org \in \bigcup_n Org_n : \text{если } \exists\ u \in \bigcup_l u_l \text{ то } g_{vo\_id}(u) = o\_id \}. \quad (8)$$

Обозначим $\bigcup_m VO_m$ все существующие ВО. Предположим, что нам необходимо найти все ВО в которых используется ресурс *r* определенной организации *Org* , или где принимает участие пользователь *u* из определенной организации *Org* с идентификатором *o_id*. Согласно (2), (3), (4):

$$\{ VO \in \bigcup_m VO_m : f_{vo\_id}(r) = o\_id \} \equiv VO|_r, \quad (9)$$

$$\{ VO \in \bigcup_m VO_m : g_{vo\_id}(u) = o\_id \} \equiv VO|_u. \quad (10)$$

Эти основные понятия используются далее для описания модели репутации для поставщиков ресурсов и пользователей.

Модель репутации для поставщиков ресурсов

Модель репутации основана на функции полезности, которая определяет уровень удовлетворенности пользователя предоставленным сервисом. Для определения функции полезности вводится вспомогательная функция согласованная между пользователем ВО и поставщиком ресурсов, которая показывает заранее оговоренное качество услуг (the service level agreement (SLA)) [16].

$$SLA : \bigcup_l u_l \times \bigcup_k r_k \times \bigcup_m VO_m \rightarrow \mathrm{R} \quad (11)$$

где R — множество действительных чисел.

Значение SLA показывает, какое качество услуг ожидает получить пользователь [16]. Метрики качество услуг (QoS), которые могут быть использованы для определения уровня удовлетворенности описаны в [17, 18]. В некоторых статьях также описываются механизмы вычисления и управления качеством услуг QoS [19].

Определение 3. Событие – это:

$$Event = T \times \bigcup_l u_l \times \bigcup_k r_k \times \bigcup_m VO_m \times \{\text{QoS name}\} \times \mathrm{R} \quad (12)$$

где T представляет временной интервал

Следовательно, событие характеризируется следующими атрибутами:

{t, u, r, vo_id, QoS, v}

где *t* показывает время, *QoS* показывает желаемый уровень качества обслуживания, а *v* реальное значение QoS определенное системой мониторинга в Grid после взаимодействия пользователя и ресурса.

Определение 4. След (Trace) представляет собой последовательность событий (12):

$$Trace = \bigcup_p Event_p = \bigcup_p \{t, u, r, vo\_id, QoS, v\}_p .$$

(13)

Перед определением репутации и функции полезности вашему вниманию предлагается три функции: первая будет характеризовать возможность сговора пользователя и ресурса с целью избегания мошенничества, вторая будет учитывать время, когда была оценена полезность, и третья будет предоставлять разные значения в зависимости от типа предоставляемого сервиса. Эти функции обеспечивают расширения для репутации функции полезности изначально представленной в [16].

Функция $h(u, r)$ будет принимать значения от 0 до 1 и показывать уровень взаимосвязи пользователя u с ресурсом r. Если между ними нет сговора $h(u, r)$ будет принимать большие значения. Например, $h(u, r)$ можно определить следующим образом:

$$h(u,r)=\begin{cases}1, \text{ if } f_{vo\_id}(r)\neq g_{vo\_id}(u)\\ \theta, \text{if } f_{vo\_id}(r)= g_{vo\_id}(u)\end{cases}, \qquad (14)$$

где θ является параметром.

Функция $z(t, t_c)$ будет показывать какие предыдущие записи о взаимодействии пользователей с ресурсами нужно принять во внимание для оценки репутации конкретного ресурса. Где t это время, а t_c –параметр. В простейшем случае $z(t, t_c)$ может быть ступенчатой функцией

$$z(t,t_c)=\begin{cases}1, t\geq t_c\\ 0, t<t_c\end{cases}. \qquad (15)$$

Функция s(type(r)) будет обеспечивать различные значения для разных типов сервисов, которые предоставляет ресурс r (функция type(r) указывает на категорию сервиса).

Теперь можно определить функцию полезности (согласно (11),(12),(14))

$$utility : \text{Event} \rightarrow \text{R},$$

$$utility(\{t,\ u,\ r,\ vo\_id,\ QoS,\ v\}) = \begin{cases} h(u,r)s(r), & \text{if } v \geq SLA(u,r,vo\_id)\\ \dfrac{v}{SLA(u,r,vo\_id)}h(u,r)s(r), & \text{if } v < SLA(u,r,vo\_id)\end{cases} \qquad (16)$$

Стоит отметить набор следов, которые используются для оценки репутации ресурса r в виртуальной организации с идентификатором vo_id до текущего времени t с:

$$\text{Trace}|_{(vo\_id,\ r,\ t)} = \left\{ \begin{array}{l} \{t', u', r', vo\_id', QoS', v'\} \in Trace : \\ r = r', vo\_id = vo\_id', t' \leq t \end{array} \right\} \qquad (17)$$

Набор значений функции полезности, полученные из $Trace|_{(vo\_id,\ r,\ t)}$, (15), (17) с учетом:

$$O_{(vo\_id,\ r,\ t)} = \{\ z(t, t_c)\cdot utility(\{t, u, r, vo\_id, QoS, v\}) \mid \{t, u, r, vo\_id, QoS, v\} \in Trace|_{(vo\_id,\ r,\ t)}\} \qquad (18)$$

Определение 5. Репутация – это математическое ожидание функции полезности *utility*() (в терминах теории вероятностей)

$$rep(vo\_id, r, t) = \text{E}[\ utility(O_{(vo\_id,\ r,\ t)})\ ] = \int utility(O_{(vo\_id,r,t)})\, p_{utility}(O_{(vo\_id,r,t)})\, dO_{(vo\_id,r,t)} \ . \qquad (19)$$

Для того чтоб различать значения функции полезности по времени будет использоваться:

$$z(t, t_c) = 1.$$

Для аппроксимации ожидания можно использовать выборочное среднее (18):

$$rep(vo\_id, r, t) = \frac{1}{\left|O_{(vo\_id,r,t)}\right|}\sum_{x\in O_{(vo\_id,r,t)}} x \qquad (20)$$

где $|\cdot|$ — количество элементов множества.

Репутация организации в виртуальной организации – это агрегация репутации всех ресурсов, которые она предоставляет для использования в виртуальной организации. Согласно (5), (15), (19):

$$rep(vo\_id, t)= \frac{1}{\left|f^{-1}_{vo\_id}(o\_id)\right|}\sum_{r\in f^{-1}_{vo\_id}(o\_id)} rep(vo\_id, r, t) . \qquad (21)$$

Репутация ресурса во всех ВО можно вычислить следующим образом (19):

$$\text{rep(r, t)} = \frac{1}{\left|\text{VO}\right|_r}\sum_{vo\_id\in \text{VO}|_r} rep(vo\_id, r, t) . \qquad (22)$$

Модель репутации для пользователей.

В модели репутации предложенной в [16] соответствующая модель для пользователя строится с использованием штрафной функции. Если пользователь предпринимает действие, которое не соответствует политике безопасности ВО или ресурса, то его штрафуют. Этот штраф используется для вычисления функции полезности для этого пользователя и соответственно репутации пользователя. Но аудит действий пользователя и проверка соответствия этих действий политике безопасности ВО или ресурса является сложной задачей, особенно с точки зрения внедрения. Должен быть четко определен критерий оценки действий пользователя и соотвествующие программные компоненты, которые будут эту оценку производить.

Мы предлагаем включить в модель репутации статистическую модель поведения пользователя, которая была ранее разработана для компьютерных сетей и в дальнейшем была модифицирована для распределенных систем, в

частности Grid систем [20, 21, 22, 23]. Эта модель основана на анализе статистических данных, которые собираются после того как пользователь выполнит действие в Grid системе. Модель была обучена и верифицирована на реальных данных собранных в инфраструктуре GILDA (https://gilda.ct.infn.it) проекта EGEE. Модель способна различать поведение разных типов пользователей и обнаруживать искусственно сгенерированные вторжения с точностью более 90%.

В частности, модель состоит из различных статистических параметров представленных следующими атрибутами:

{*S*, *ET*, *CPU*, *WT*, *CW*, *ES*, *CT*, *STD*, *RAM*, *VM*, *VO*, *RB*}

где *S* – сайт на котором выполнялась задача, *ET(*execution target*)* – ресурс сайта, на котором выполнялась задача, *CPU* (CPU time) – время работы процессора ресурса при выполнении задачи, *WT* (wall time) – полное время выполнения задачи, *CW (*CPUWall = CPU/W) – отношение времени работы процессора к общему времени выполнения задачи, ES (exit status) – статус завершения задачи (успешное завершение или с ошибкой), *CT* (creation time) – время создания (отправки) задачи в Grid систему, *STD* (start time difference) – разница между временем начала выполнения задачи на выбранном ресурсе Grid системы и временем отправки задачи в Grid систему, *RAM (*RAM used) – использованная оперативная память, *VM* (virtual memory used) – использованная виртуальная память, *VO* – принадлежность к ВО, *RB* (resource broker) – брокер ресурсов, который был использован для распределения задачи.

Этот набор параметров используется для выявления аномальных сигнатур в действиях пользователя для поднятия тревоги в системе. Например, такие сигнатуры могут включать ситуации, когда задача выполняется значительное количество времени, либо когда процессор ресурса загружен на 100% [24]. С целью выявления таких сигнатур в данных которые были записаны в течении мониторинга действий пользователя мы используем нейронные сети [25]. Нейронная сеть обучается для каждого пользователя так чтоб отличать нормальное и аномальное поведение пользователя. Когда нейронная сеть обучена, то целевой выход – это значение от 0 (данные соответствующие аномальному поведению пользователя) до 1 (данные соответствующие нормальному поведению пользователя). Для того чтоб представить оба случая (нормальное и аномальное поведение) в обучающей выборке были использованы данные Grid системы мониторинга: данные о предыдущем поведении пользователя представляют нормальное поведение, а данные о поведении других пользователей и искусственно сгенерированные представляют аномальное поведение пользователя. Искусственные данные могут быть сгенерированные с использованием генеративных моделей (generative models) и встраиваться в обучающую выборку.

Таким образом, мы предлагаем использовать выход модели данной модели для оценки репутации пользователя в ВО. Такая модель пользователя будет характерной для ВО в зависимости от ее целей и типов задач, которые в ней исполняются. Например, Во может быть ориентирована на приложения, которым требуется выполнение большого количества задач со сравнительно небольшим количеством данных, которые обрабатываются в одной задаче. В таких ВО задачи, которые потребляют практически вю виртуальную память ресурса будут считаться аномальными. С другой стороны, другие ВО могут быть ориентированны на приложения где одна задача состоит из из многочисленных элементарных подзадач, каждая из которых обрабатывает большое количество данных (например науки о Земле или обработка спутниковых данных [26, 27]).

Приведем формальное описание модели репутации для пользователей основываясь на статистической модели поведения пользователя как было приведено выше для поставщиков ресурсов.

Определение 6. Событие для пользователя – это:

$$Event = \{t, u, r, vo\_id, x\}. \qquad (23)$$

где $x = (S, ET, CPU, WT, CW, ES, CT, STD, RAM, VM, VO, RB)$.

Определение 7. След - это:

$$Trace = \bigcup_p Event_p = \bigcup_p \{t, u, r, vo\_id, \mathbf{x}\}_p \qquad (24)$$

Аналог функции полезности (*utility*() function):

$$utility : \text{Event} \rightarrow R,$$
$$utility(\{t, u, r, vo\_id, \mathbf{x}\}) = SMUB_{(u,\ vo\_id)}(\mathbf{x}), \qquad (25)$$

где $SMUB_{(u,\ vo\_id)}(\mathbf{x})$ это выход статистической модели поведения пользователя.

Следует отметить, что в общем случае под функцией полезности для пользователей подразумевается, что можно использовать другие модели поведения пользователей (т.к. [28, 29]) или комбинацию нескольких моделей.

В нашем случае, согласно (24), преобразования в статистической модели поведения пользователя производятся нейронной сетью и являются специфичными для пользователя и ВО. Отметим, что набор следов которые используются для оценки репутации пользователя u в ВО с идентификатором vo_id в настоящее

время t:

$$Trace|_{(vo\_id,\,u,\,t)} = \{\{t', u', r', vo\_id', \mathbf{x}'\} \in Trace : u = u', vo\_id = vo\_id', t' \le t\}. \quad (26)$$

Значения функции полезности *utility()* полученные из следов $Trace|_{(vo\_id,\,u,\,t)}$, согласно (15), (25), (26) с

$$O_{(vo\_id,\,u,\,t)} = \{ z(t, t_c) \cdot utility(\{t, u, r, vo\_id, \mathbf{x}\}) \mid \{t, u, r, vo\_id, \mathbf{x}\} \in Trace|_{(vo\_id,\,u,\,t)}\} \quad (27)$$

Определение 8. Репутация – это ожидание функции полезности (25),(27)

$$rep(vo\_id,\ u,\ t) = \mathrm{E}[\ utility(O_{(vo\_id,\ u,\ t)})\ ] = \int utility(O_{(vo\_id,u,t)})\, p_{utility}(O_{(vo\_id,u,t)})\, dO_{(vo\_id,u,t)} \ . \quad (28)$$

Для аппроксимации ожидания можно использовать выборочное среднее

$$rep(vo\_id, u, t) = \frac{1}{\left|O_{(vo\_id,u,t)}\right|} \sum_{x \in O_{(vo\_id,u,t)}} x \ . \quad (29)$$

Репутация организации в ВО (с точки зрения пользователя) это агрегация репутации всех пользователей которые представляют эту организацию в данной ВО. Согласно (6), (29)

$$rep(vo\_id, t) = \frac{1}{\left|g^{-1}_{vo\_id}(o\_id)\right|} \sum_{r \in g^{-1}_{vo\_id}(o\_id)} rep(vo\_id, u, t) . \quad (30)$$

Согласно (2), (10), (29) репутацию пользователя во всех ВО можно оценить:

$$rep(r, t) = \frac{1}{\left|\mathrm{VO}\right|_u} \sum_{vo\_id \in \mathrm{VO}|_u} rep(vo\_id, u, t) . \quad (31)$$

Результаты экспериментов

Для проверки адекватности и эффективности построенной модели доверия, была проведена серия экспериментов. Для этого использовались реальные данные о параметрах запускаемых задач в Grid-системе EGEE. Эти данные предоставляются в рамках проекта Grid Observatory (www.grid-observatory.org), который является частью проекта EGEE-III EU INFSO-RI-222667.

В рамках данной статьи использовались данные системы Real Time Monitor (RTM), которая предоставляет информацию о 37 параметрах каждой задачи, в том числе время запуска, время ожидания, процессорное время и т.д. Всего в работе использовалась статистическая информация о 35000 задач, которые исполнялись в период с 30.06.2009 по 13.07.2009.

В качестве основной метрики качества обслуживания использовалось время выполнения задачи на ресурсе Grid-системы. Значение уровня SLA (Service Level Agreement) (11), которое использовалось для оценки функции полезности (16), вычислялось как среднее значение сложности задач, деленное на среднее значение производительности ресурса. Для проверки масштабируемости предложенного подхода, эксперименты проводились для различного набора задач и для различных конфигураций Grid-системы (с разных количеством ресурсов и разным значением производительности). Кроме того, были проведены эксперименты по использованию репутации для распределения задач между ресурсами Grid-системы. Информация о репутации использовалась в планировщике задач следующим образом: количество задач в очереди ресурса было прямо пропорционально его репутации. Рассмотрим полученные результаты более подробно.

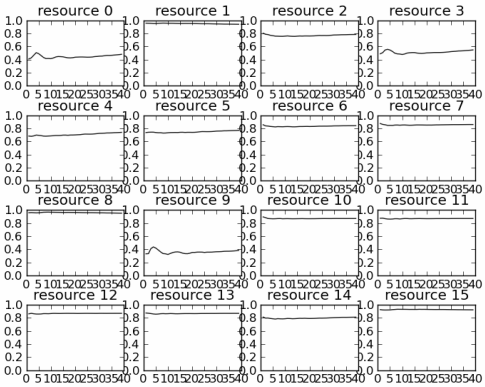

Рисунок 2 – Изменение репутации ресурсов во времени

На рис. 2 представлено изменение во времени сглаженных значения репутации ресурсов.

Пример изменения репутации ресурса с уменьшением масштаба времени представлен на рис. 3.

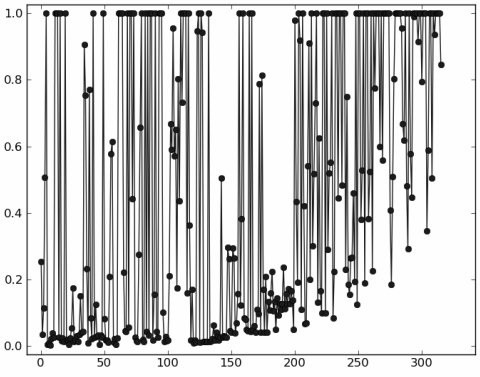

Рисунок 3 – Изменение репутации ресурса во времени

Рассмотренный метод планирования с учетом репутации сравнивался с алгоритмом циклического планирования. Зависимость времени выполнения задач при разных планировщиках для

разных наборов задач представлена на рис. 4.

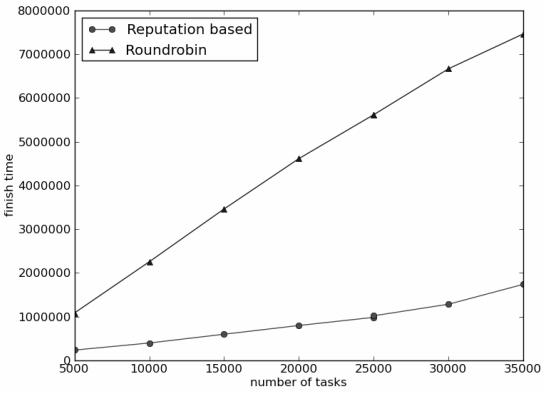

Рисунок 4 – Сравнение планировщика, который учитывает репутацию (reputation based), с алгоритмом циклического планирования (roundrobin)

Из рис. 4 видно, что использование планировщика, который учитывает репутацию, позволяет уменьшить время выполнения задач в среднем более чем на 80%.

Кроме того, были построены графики зависимости репутации ресурса от его производительности (рис. 5), используя планировщик с репутацией. В результате получена нелинейная зависимость, причем значение репутации ресурса возрастает с ростом его производительности. Это объясняется тем, что в качестве основной метрики качества обслуживания использовалась время выполнения заданий, которое зависит от производительности ресурса, и отсутствием атак на систему оценки репутации. Учет других метрик качества обслуживания, моделирование атак на систему репутации, а также усовершенствование планировщика, учитывающего репутацию ресурса, является целью дальнейших исследований.

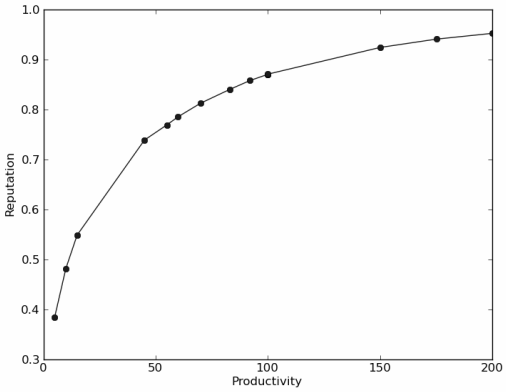

Рисунок 5 – Изменение репутации в зависимости от производительности (значение производительности ресурса представлено в условных единицах)

Выводы

В данной статье рассмотрены основные подходы к построению моделей доверия для Grid-систем на основании репутации. Предложено расширение существующей модели репутации для виртуальных организаций в Grid-системах, основанной на вычислении функции полезности и параметров качества обслуживания. Для проверки адекватности и эффективности построенной модели доверия, была проведена серия экспериментов, используя реальные данные о параметрах запускаемых задач в Grid-системе EGEE. Проведено сравнение алгоритма планирования задач в Grid-системе с учетом репутации и циклического планирования. Результаты экспериментов показали, что для разных наборов задач учет репутации позволяет уменьшить время выполнения задач более чем на 80%.

Литература

1. Foster I., Kesselman C., Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations // International Journal of Supercomputing Applications, 15(3), 2001. – p. 200-222,.
2. Castelfranchi C., Falcone R., Sadighi B., Tain Y.-H. Guest Editorial. Applied Artificial Intelligence, 14(9), 2000, Taylor & Frances,.
3. Waidner M.. Ercim News, Special Theme: Information Security. No 49, 2002.
4. Nixon P., Terzis S. First International Conference on Trust Management // Lecture Notes in Computer Science, vol. 2692, Springer, 2003.
5. Jensen C.D., Poslad S., Dimitrakos T. Second International Conference on Trust Management // Lecture Notes in Computer Science, vol. 2995, Springer, 2004.
6. Hermann P., Issarny V., Shue S. Third International Conference on Trust Management // Lecture Notes in Computer Science, vol. 3477, Springer, 2005.
7. Grandison T., Sloman M. A Survey of Trust in Internet Applications // IEEE Communications Survey and Tutorials, 3, 2000.
8. McKnight D.H., Chervany N.L. The Meaning of Trust // Technical Report MISRC Working Paper Series 96-04, University of Minnesota. Management Information Systems Research Center, 1996.
9. Gambetta D. Can We Trust Trust? In D. Gambetta (editor). Trust: Making and Breaking Cooperative Relations. Department of Sociology, Univ. of Oxford, 1988.

10. Josang A., Ismail R., Boyd C. A Survey of Trust and Reputation Systems for Online Service Provision // Decision Support Systems, 43(2), 2007. – p. 618-644,.
11. Rasmusson L., Janssen S. Simulated Social Control for Secure Internet Commerce // In C. Meadows. Proceedings of the 1996 New Security ParadigmsWorkshop. ACM.
12. CoreGrid. D.ia.03 survey material on trust and security. Technical Report D.IA.03, CoreGrid, October2005. http://www.coregrid.net/mambo/images/stories/IntegrationActivities/TrustandSecurity/d.ia.03.pdf.
13. Abdul-Rahman A., Hailes S. Supporting trust in virtual communities // In HICSS '00: Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 6, page 6007, Washington, DC, USA, 2000. IEEE Computer Society.
14. Kerschbaum F., et al. A trust-based reputation service for virtual organization formation. In Proceedings of the 4th International Conference on Trust Management, vol. 3986 of Lecture Notes in Computer Science, pp. 193–205. Springer, 2006.
15. Luke T.W.T., Jennings N.R., Rogers, Luck M. A Hierarchical Bayesian Trust Model based on Reputation and Group Behaviour // 6th European Workshop on Multi-Agent Systems, 18th-19th December, 2008, Bath, UK.
16. Arenas A.E., Aziz B., Silaghi G.C. Reputation Management in Grid-Based Virtual Organisations // Proc. International Conference on Security and Cryptography (SECRYPT 2008), Porto, Portugal, 26-29 Jul 2008, INSTICC.
17. Menasce D.A., Casalicchio E. Quality of service aspects and metrics in Grid computing // In: Proc. 2004 Computer Measurement Group Conference, Las Vegas, USA, 2004.
18. Hong-Linh T., Samborski R., Fahringer T. Towards a Framework for Monitoring and Analyzing QoS Metrics of Grid Services // In: Proc. Second IEEE Int Conf on e-Science and Grid Computing (e-Science'06), 2006.
19. Al-Ali R., von Laszewski G., Amin K., Hategan M., Rana O., Walker D., Zaluzec N. QoS Support for High-Performance Scientific Grid Applications // In: Proc. IEEE International Symposium on Cluster Computing and the Grid 2004. (CCGrid 2004). – p. 134–143.
20. Shelestov A., Skakun S., Kussul O. Intelligent Model of User Behavior in Distributed Systems // International Journal on Information Theory and Applications, Volume 15, Number 1, 2008. – p. 70-76.
21. Shelestov A., Skakun S., Kussul O. Complex Neural Network Model of User Behavior in Distributed Systems // Proc. of XIII-th International Conference Knowledge-Dialogue-Solutions. - 2007. - Varna, Bulgaria. - P. 42-49.
22. Куссуль Н.Н. Реализация нейросетевой модели пользователей компьютерных систем на основе агентной технологии / Н.Н. Куссуль, С.В. Скакун, А.Г. Лобунец // Проблемы управления и информатики, № 2, 2005. — С. 93–102.
23. Kussul N. Skakun S. Neural Network Approach for User Activity Monitoring in Computer Networks. In: Proc. of the International Joint Conference on Neural Networks. — Budapest (Hungary), Vol. 2, 2004. – p. 1557-1562.
24. Chakrabarti A. Grid Computing Security. Springer-Verlag Berlin Heidelberg, 2007.
25. Haykin S. Neural Networks: A Comprehensive Foundation. Upper Saddle River, New Jersey, Prentice Hall, 1999.
26. Shelestov A., Kussul N., Skakun S. Grid Technologies in Monitoring Systems Based on Satellite Data // J. of Automation and Inf. Sci., 38(3), 2006. – p. 69–80.
27. Fusco L., Cossu R., Retscher C. Open Grid Services for Envisat and Earth Observation Applications // In: Plaza AJ, Chang C-I (ed) High performance computing in remote sensing, 1st edn. Taylor & Francis Group, New York, 2007. – p. 237-280.
28. Oh S.H., Lee W.S. An anomaly intrusion detection method by clustering normal user behaviour // Computers & Security, 22(7), 2003. – p. 596–612.
29. Wang W., Guan X., Zhang X., Yang L. Profiling program behavior for anomaly intrusion detection based on the transition and frequency property of computer audit data // Computers & Security, 25(7), 2006. – p. 539–550.

Поступила в редакцию 30.03.2010