

УДК 004.021

**ФИЗИЧЕСКОЕ МОДЕЛИРОВАНИЕ МЕТОДОВ ОБЕСПЕЧЕНИЯ
ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ НА ОСНОВЕ ДОВЕРИЯ И
РЕПУТАЦИИ**

Мельников Т.Ю. (Университет ИТМО)

Научный руководитель – кандидат технических наук, научный сотрудник ФБИТ

Виксин И.И.
(Университет ИТМО)

В докладе описывается алгоритм, позволяющий элементам мультиагентной системы самостоятельно определять достоверность информации, получаемой друг от друга. Применение данного алгоритма повышает как автономность системы, так и защищенность информации внутри нее.

Введение. На данный момент существует проблема оценки качества и достоверности информации, которой обмениваются агенты или которую они получают извне. Разработана концепция Data Quality, предлагающая общий подход к решению подобных проблем через оценку качества данных, но не дающая конкретного решения. Другими словами, для каждой ситуации нужно продумывать отдельный вариант применения этого подхода.

Основная часть. Рассмотрим частный способ реализации концепции Data Quality. Пусть есть мультиагентная система, состоящая из беспилотных транспортных средств. Агенты обмениваются между собой информацией об окружающей среде или друг о друге по общему каналу связи. Требуется добавить им возможность проверять эту информацию без внешнего вмешательства в систему.

Введем три величины, каждая из которых лежит в диапазоне $[0;1]$:

- Истинность
- Репутация
- Доверие

Истинность представляет из себя функцию от входных данных и отражает оценку достоверности этих данных самим агентом.

Репутация – это функция от истинности, показывающая, насколько достоверно агент оценивает поступающую информацию. Каждый агент имеет свое значение репутации в системе, которое учитывается при принятии решения. Чем выше данное значение, тем больше вероятность того, что оценка агента соответствует действительности. Слишком низкое значение репутации ведет к тому, что агента исключают из системы, и его оценки перестают учитываться.

Доверие является функцией от предыдущих двух значений и служит заключением относительно поставленного вопроса. Если доверие больше некоторого доверительного значения, определенного заранее, то агенты считают, что информации можно доверять, иначе суждение признается ложным.

Когда требуется проверить информацию, поступающую от одного из агентов, выполняется следующая последовательность действий:

1. Каждый агент оценивает поступившую информацию, отправляя в общий канал значение истинности: 1, если информация считается истинной и 0, если ложной.
2. Высчитывается среднее значение истинности, исходя из всех полученных от агентов оценок.

3. Происходит перерасчет репутации агента, отправившего изначально информацию, на основе полученного значения истинности и предыдущих значений его репутации.
4. Результат нормируется по времени, чтобы не выходить за границы диапазона [0;1]
5. Считается доверие на основе полученных ранее значений истинности и репутации и сравнивается с доверительным значением, на основе чего делается вывод, можно ли доверять полученной информации.

Выводы. Для оценки эффективности данного подхода в лаборатории была создана мультиагентная система, состоящая из беспилотных транспортных средств. Агенты обменивались информацией о своем местоположении, которая периодически нуждалась в проверке, если агент не обновлял ее некоторое время. Изначально время ожидания новой информации было фиксировано и составляло 120 секунд (значение было выявлено экспериментально, как доверительное). Затем был использован подход на основе доверия и репутации, причем время ожидания было привязано к значению репутации агента. Таким образом, чем ниже репутация агента, тем меньшее время от него ждут новой информации. Такой подход позволил снизить время ожидания примерно до 41-43 секунд. Кроме того, в таком случае агенты вынуждены чаще перепроверять информацию друг о друге, что повышает надежность работы всей системы. Можно сделать вывод, что предложенный метод представляет практическую ценность и способен заметно улучшить работу многих мультиагентных систем.

Мельников Т.Ю. (автор)

Подпись

Виксин И.И. (научный руководитель)

Подпись