ОБ АВТОРАХ

Верхотуров Михаил Александрович, проф., каф. выч. матем. и киберн. Дипл. инж.-системотех. по АСУ (УАИ, 1983). Д-р техн. наук по АСУ (УГАТУ, 2000). Иссл. в обл. оптимизации размещения двух- и трехмерн. геометрич. объектов.

Садретдинова Нелли Маратовна, зав. сектором в «ЮНГ-НТЦ Уфа». Дипл. мат.-экон. (УГАТУ, 2001). Канд. техн. наук по мат. и прогр. обеспеч. выч. машин, комплексов и компьют. сетей (УГАТУ, 2004). Иссл. в обл. интернет/интранет-технологий.

УДК 004.4'2

Г. А. МАКЕЕВ

СВОЙСТВО УСТОЙЧИВОСТИ ДЛЯ СИСТЕМ РАСПРЕДЕЛЕННОЙ СОВМЕСТНОЙ ФИЛЬТРАЦИИ ИНФОРМАЦИИ

Рассматривается теоретический подход к анализу системы совместной фильтрации. Для этого формализуются критерии управляемости и устойчивости по пользователям и сообщениям. Предлагаются различные виды функций транзитивного замыкания и агрегации сообщений, производится их формальный анализ, позволяющий выбрать наилучшую из них по соответствию формализованным критериям. *Системы совместной фильтрации; устойчивость; формализация*

Интеграция новых информационных технологий с глобальными компьютерными сетями, с одной стороны, увеличивает объем данных, доступных и подлежащих обработке, а с другой стороны, предоставляет возможность организации распределенных вычислений, в том числе и для распределенной обработки данных.

К этой области, в частности, относятся получающие все более широкое распространение и применение системы рекомендаций (Recommender Systems) [6] и системы совместной фильтрации информации (Collaborative Filtering Systems) [8].

Системы совместной фильтрации требуют от пользователя сначала предоставить некоторый набор своих рекомендаций, выявляя тем самым его предпочтения. Основной целью системы является определение на основе рекомендаций тех пользователей, предпочтения которых схожи с предпочтениями данного пользователя, и формирование рекомендаций для него на основе рекомендаций найденных пользователей.

Пользователи фиксируют свои предпочтения, оценивая фильтруемые элементы (сообщения, книги, звукозаписи, и т. д.). Традиционная система совместной фильтрации, например, такая как GroupLens [7], ищет пользователей со схожими предпочтениями среди всех своих пользователей. Далее система предлагает пользователю в качестве результата совместной фильтрации те элементы, которые были высоко оценены «похожими» пользователями. Таким образом, из всего множества фильтруемых элементов пользователь получает только некоторый набор, потенциально более ценный для него.

По мере распространения систем совместной фильтрации информации, по мере роста числа пользователей, вовлеченных в них, и по мере увеличения объема обрабатываемой ими информации все больше начнут проявляться качественные аспекты функционирования, в частности, ограничения таких систем, и все более важными будут вопросы их организации для наиболее эффективного функционирования.

Одним из таких вопросов является устойчивость системы совместной фильтрации к разрушающим воздействиям со стороны «плохих» пользователей. Предлагаемый в данной работе формальный аппарат доказательства тех или иных аспектов устойчивости систем совместной фильтрации позволит строить системы совместной с качественно более высоким уровнем защищенности от злонамеренных действий других пользователей.

Вопросы безопасности в существующих системах совместной фильтрации. Общее описание известных подходов к системам рекомендаций было приведено автором в работе [1].

Исследования проводились в рамках Федеральной целевой программы «Интеграция науки и высшего образования РФ на 2002–2006 гг.» по проекту «Фундаментальные исследования и новые технологии проектирования сложных технических систем» и частично поддержаны грантом РФФИ 03-07-90242 «Интернет-комплекс поддержки выполнения проектов фундаментальных исследований сложных систем с применением интеллектуальных технологий на базе экспертных систем» (2003–2005 гг.).

Большинство существующих систем рекомендаций являются централизованными, что приводит к целому ряду сложноразрешимых проблем. Во-первых, при таком подходе пользователь не может управлять процессом построения рекомендации для него. Во-вторых, такая система становится неустойчивой, зависимой от функционирования центрального сервера. В-третьих, она небезопасна с точки зрения хранения конфиденциальной информации на удаленной от пользователя машине. Критика централизованной организации системы рекомендаций представлена, например, в работе [2]. Там же рассматривается децентрализованный подход к построению системы рекомендаций и метод вычисления некоторым сообществом пользователей агрегированных рекомендаций, при котором индивидуальные конфиденциальные рекомендации не открываются.

Во всех разработанных к настоящему времени системах практически отсутствует контроль пользователя за созданием рекомендаций, что влечет за собой отказ пользователей от системы в случае, если она начинает давать плохие рекомендации, из-за невозможности на нее влиять [3].

Неуправляемость пользователем процесса построения рекомендаций в большинстве существующих систем совместной фильтрации приводит к еще одной критической, на наш взгляд, проблеме — уязвимости подобных систем к пользователям, намеренно нарушающим нормальную работу системы. Раз репутация пользователя зависит от его оценок со стороны всех пользователей, то достаточно большой группе пользователей ничего не мешает сильно увеличить или уменьшить репутацию некоторого пользователя и повлиять, таким образом, на всех остальных. Но еще более важно то, что предлагаемые конкретному пользователю рекомендации слабо зависят от его участия, зато сильно зависят от поведения других пользователей. В P2P-системах обмена файлами [5], например, большая группа пользователей может высоко оценить некоторый злокачественный файл, и их оценки повлияют на всех остальных пользователей.

Следует отметить, что в большинстве подходов к построению систем рекомендаций и систем совместной фильтрации, в частности, оценка безопасности и устойчивости системы к разрушающим воздействиям проводится на эвристическом уровне, что не позволяет формально оценить возможное поведение системы при тех или иных атаках.

Структура системы распределенной совместной фильтрации. В предлагаемой распределенной системе совместной фильтрации [4] каждый пользователь системы U_i (здесь и далее $i \in [1..N^U]$) является владельцем узла peer-to-peer сети (в случае идеальной сетевой инфраструктуры, подразумевающей постоянное включение узлов в сеть и отсутствие межсетевых экранов).

Каждый пользователь U_i размещает у себя на узле множество сообщений $M_i = \{m_k : k \in [1..N_i^m]\}$, подписывая их своей парой ключей шифрования, так что любой пользователь может убедиться в авторстве сообщения. Содержание некоторого сообщения m во множестве M_i означает, что пользователь U_i рекомендует это сообщение (т. е. считает высокой его ценность) (рис. 1).

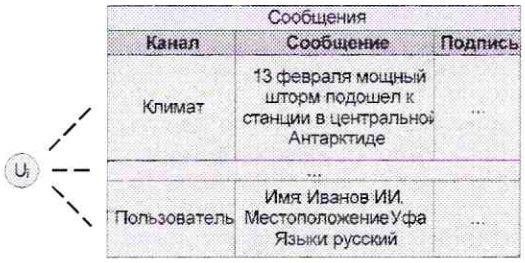

Рис. 1. Управление сообщениями

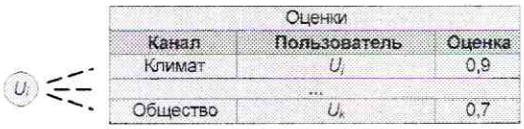

Рис. 2. Множество оценок пользователя

Кроме множества сообщений, хранимых пользователем на своем узле, пользователь управляет множеством своих оценок других пользователей $R_i = \{(U_j, v_{ij}) : j \in [1..N^U]\}$. Только сам пользователь может добавлять и удалять свои оценки других пользователей, другие пользователи не могут влиять на то, как их оценивает данный пользователь (рис. 2).

Значение v_{ij}, непосредственно задаваемое пользователем, отражает его осознаваемую оценку другого пользователя, ценность его предпочтений и стремление использовать результаты его труда.

Однако пользователь может оценивать довольно ограниченный круг других людей. Построить расширенное множество пользователей R_i^*, включаемых в фильтрацию, позволяет механизм транзитивности оценок пользователей. Проще говоря, для трех пользователей U_i, U_j, U_k, если U_i в какой-то мере «доверяет» U_j, а U_j «доверяет» U_k, то U_i может в какой-то мере «доверять» и U_k. Основной задачей здесь является выбор функции $TRF(U_i, U_j, U_k)$, используемой для вычисления транзитивно замыкаемой оценки, потому что от этой функции, в том числе, непосредственно зависит функционирование системы.

При построении результата совместной фильтрации собираются все сообщения выбранных пользователей, и каждому сообщению присваивается некий ранг, отражающий то, как часто это сообщение встречалось у выбранных пользователей и как велики были оценки тех пользователей, у кого оно встретилось. Ранг сообщения вычисляется с помощью функции вычисления рейтинга сообщений $AMF(m, R_i^*)$, которая определяет то, каким образом оценка конкретного сообщения M вычисляется по расширенному множеству оценок R_i^*.

Формализация свойства устойчивости. Выбор оптимального вида функции транзитивного замыкания $TRF(U_i, U_j, U_k)$ и функции $AMF(m, R_i^*)$ является самой важной задачей организации системы. Для того чтобы сравнивать разные функции между собой и выявлять более подходящую, необходимо иметь объективные критерии оценки применимости этих функций для системы.

Устойчивость процесса совместной фильтрации для некоторого пользователя заключается в невозможности других пользователей нарушать избирательность результата совместной фильтрации для данного пользователя. Под нарушением этого процесса можно понимать разнообразные ситуации и, соответственно, получать различные критерии, характеризующие те или иные аспекты устойчивости.

В частности, одной из ситуаций, нарушающих нормальный процесс совместной фильтрации, можно считать ту, когда некоторое подмножество пользователей $U' \subset U$ может добиться того, что оценка пользователем U_i пользователя U_k возрастает, приближаясь к 1. Это отражает нежелательное влияние сторонних пользователей на процесс вычисления оценки для некоторого пользователя.

В терминах предложенной модели этот аспект устойчивости по пользователям можно формализовать в виде функции f:

$$\begin{gathered}\forall U_i \text{ имеет место } f(R_i^L) = \max(v_{ij}) \\ \text{по всем } (U_j, v_{ij}, L+1) \in R_i^{L+1},\end{gathered} \tag{1}$$

описывающей зависимость максимальной вычисленной оценки пользователей, включаемых в расширенное множество оценок следующего уровня транзитивности, от оценок пользователей предыдущего уровня транзитивности. Этот критерий описывает поведение оценок при расширении множества пользователей, включаемых в фильтрацию.

Аналогично можно ввести понятие устойчивости по сообщениям, отражающее возможное влияние других пользователей на вычисляемый ранг некоторого сообщения в ранжированном результате фильтрации. Формализовать это влияние можно в виде функции

$$\begin{gathered}\forall U_i \text{ имеет место} \\ f(M'_i, R'_i, M'_j, R'_j) = \max(AMF(m, R_i^*)), \\ \text{по всем сообщениям } M \in \bigcup_{U_j} M_j \text{ по всем } U_j, \\ \text{таким, что } (U_j, v_{ij}, L_{ij}) \in R_i^*,\end{gathered} \tag{2}$$

которая описывает зависимость максимального ранга некоторого сообщения в результате фильтрации для пользователя U_i от параметров, изменяемых некоторым пользователем U_j.

Выбор вида функции транзитивности оценок $TRF(u_i, u_j, u_k)$. Анализ выполнения свойства устойчивости. Функция транзитивности $TRF(U_i, U_j, U_k)$ используется при построении расширенного множества оценок, включая в совместную обработку тех пользователей, которые непосредственно не оценены пользователем.

По определению $(U_k, TRF(U_i, U_j, U_k), L_{ij} + 1) \in R_i^*$, если $(U_j, v_{ij}, L_{ij}) \in R_i^*$ и $(U_k, v_{jk}) \in R_j$, $L_{ij} < L$, где R_i^* есть расширенное множество оценок пользователя U_i, а R_j есть обычное множество оценок пользователя U_j, т. е. функция TRF используется при вычислении оценок для пользователей следующего уровня транзитивности (рис. 3).

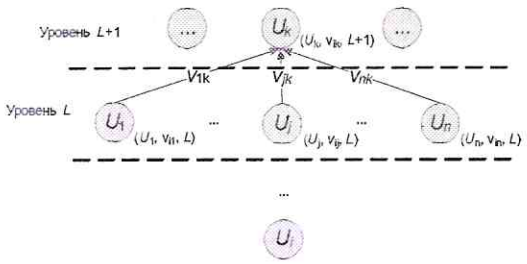

Рис. 3. Вычисление $(L+1)$-го уровня через L-й

Было предложено два альтернативных метода вычисления транзитивных оценок:

$$\begin{gathered}\frac{\sum\limits_j (v_{ij} * v_{jk})}{\sum\limits_j (v_{ij})} \text{ по всем } j \text{ таким,} \\ \text{что } (U_j, v_{ij}, L) \in R_i^L;\end{gathered} \tag{3}$$

$$\begin{gathered}\frac{\sum\limits_j ((v_{ij})^2 * v_{jk})}{\sum\limits_j (v_{ij})} \text{ по всем } j \text{ таким,} \\ \text{что } (U_j, v_{ij}, L-1) \in R_i^*.\end{gathered} \tag{4}$$

Используя приведенные выше критерии, удалось показать, что для первого метода вычисления транзитивной оценки $f_1(R_i^L) = \max\limits_{U_j}(v_{ij}) = 1$, а $f_2(R_i^L) = \max\limits_{U_j}(v_{ij})$ по всем $(U_j, v_{ij}, L) \in R_i^L$. Это означает, что в случае использования первого метода некоторые пользователи $U_1 \ldots U_N$ могут, сговорившись, добиться того, что оценка пользователем U_i произвольного пользователя U_k будет равна 1, т. е. максимально высокой — чего пользователь U_i, возможно, вовсе и не хотел. В случае использования второго метода вычисляемая оценка v_{ik} пользователей уровня транзитивности $L+1$ не может быть больше максимальной оценки пользователей уровня L, т. е. пользователи $U_1 \ldots U_N$, даже сговорившись, не смогут добиться, что оценка некоторого U_k превысит их собственные оценки.

Таким образом, $f_2(R_i^L) = \max\limits_{U_j}(v_{ij}) \leqslant 1 = f_1(R_i^L)$, и, формально говоря, с точки зрения этого аспекта устойчивости, второй предлагаемый метод вычисления транзитивных оценок количественно лучше первого.

Выбор вида функции агрегации сообщений $AMF(m, R_i^*)$. Анализ выполнения свойства устойчивости. Функция вычисления ранга сообщения $AMF(m, R_i^*)$ используется при построении результата совместной фильтрации как ранжированного множества сообщений.

Неранжированый результат совместной фильтрации определяется как объединение множеств сообщений, рекомендуемых пользователями, попавшими в расширенное множество оценок пользователя, т. е. $M_i^* = \bigcup_{U_j} (M_j)$ по всем U_j, таким, что $(U_j, v_{ij}, L_{ij}) \in R_i^*$.

Для агрегации и оценки собранных сообщений строится ранжированный результат совместной фильтрации MR_i^* для пользователя U_i путем присваивания каждому сообщению из неранжированного результата фильтрации некоторого рейтинга. Согласно (2.10), MR_i^* есть множество пар $(m, AMF(m, R_i^*))$, где $m \in M_i^*$, а $AMF(m, R_i^*)$ есть некоторая числовая функция, определяющая то, каким образом оценка конкретного сообщения m вычисляется по расширенному множеству оценок R_i^*.

Рассмотрим одного пользователя уровня $L-1$ и нескольких пользователей уровня L (рис. 4). Предположим также, что все эти пользователи содержат сообщение m в своем множестве сообщений.

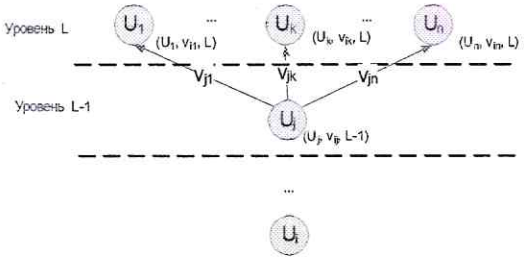

Рис. 4

Было предложено два альтернативных метода вычисления ранга сообщений:

$$AMF(m, R_i^*) = \frac{s}{S},\ \text{где } s = \sum_k v_{ik} \text{ по всем } k \text{ таким, что } (U_k, v_{ik}, L) \in R_i^* \text{ и } m \in M_k^*,\ S = \sum_k v_{ik} \text{ по всем } k \text{ таким, что } (U_k, v_{ik}, L) \in R_i^*; \quad (5)$$

$$AMF(m, R_i^*) = \frac{s}{S},\ \text{где } s = \sum_k (v_{ik})^2 \text{ по всем } k \text{ таким, что } (U_k, v_{ik}, L) \in R_i^* \text{ и } m \in M_k^*,\ S = \sum_k v_{ik} \text{ по всем } k \text{ таким,что } (U_k, v_{ik}, L) \in R_i^*. \quad (6)$$

Используя приведенные выше критерии устойчивости по сообщениям, удалось показать, что для первого метода вычисления ранга сообщения $f_1(M'_i, R'_i, M'_j, R'_j) = \max(AMF(m, R_i^*)) = 1$. Этот результат говорит о том, что любой пользователь, попавший в расширенное множество оценок пользователя U_j и находящийся не на последнем уровне транзитивности, может добиться того, что ранг сообщения m для пользователя U_i будет больше ранга любого другого сообщения. Это также говорит о неустойчивости системы, о ее открытости для потенциальных злоупотреблений со стороны пользователей.

Для второго метода вычисления ранга сообщения $f_2(M'_i, R'_i, M'_j, R'_j) = \max(AMF(m, R_i^*)) = v_{ij}$ для всех $(U_j, v_{ij}, L) \in R'^*_i$. Этот результат свидетельствует о том, что любой пользователь U_j, попавший в расширенное множество оценок пользователя U_i и находящийся не на последнем уровне транзитивности, не может добиться того, что ранг сообщения в ранжированном результате совместной фильтрации для пользователя U_i будет превышать вычисленную оценку самого пользователя U_j, которая от него самого не зависит.

Таким образом, $f_2 = v_{ij} \leqslant 1 = f_1$, и второй метод лучше первого. С практической точки зрения это означает, что чем ниже оценка у пользователя, тем меньшее влияние он может оказать на ранг сообщения.

ЗАКЛЮЧЕНИЕ

Таким образом, рассмотренный формальный подход к анализу свойств системы распределенной совместной фильтрации информации позволяет доказывать выполнение различных свойств системы при разных параметрах организации. Так, из предложенных вариантов функций формальный анализ позволил не только выбрать более хорошую с точки зрения устойчивости, но и выявить неочевидные возможности атаки злонамеренных пользователей.

СПИСОК ЛИТЕРАТУРЫ

1. **Макеев, Г. А.** Системы рекомендаций. Анализ проблемы и известных подходов к решению / Г. А. Макеев // Автоматизированные системы обработки информации и управления : сб. тр. мк. асп. Уфа : УГАТУ, 2006. Т. 1. С. 6–11.
2. **Canny, J.** Collaborative filtering with privacy / J. Canny // IEEE Conf. on Security.
3. **Massa, P.** Trust-aware Decentralized Recommender Systems: PhD research proposal / Dept. of Information and Communication Technology. Univ. of Trento, Italy, 2003.
4. **Макеев, Г. А.** Совместная фильтрация информации, устойчивая к внешним воздействиям / Г. А. Макеев, Н. И. Юсупова, Д. В. Попов // 7-я Междунар. конф. CSIT'2005. Уфа–Ассы, 2005. Т. 3. С. 209–215. (На англ. яз.).
5. **Cornelli, F.** Implementing a reputation-aware gnutella servent / F. Cornelli [et al] // Int. Workshop on Peer-to-Peer Computing. May 2002.

6. **Resnick, P.** Recommender systems / P. Resnick, H. R. Varian // Communications of the ACM. 1997. Vol. 40 (3). P. 56–58.
7. **Resnick, P.** GroupLens: An open architecture for collaborative filtering of netnews / P. Resnick [et al] // Proc. of CSCW'94. Chapel Hill NC: ACM Press, October 1994. P. 175–186.
8. **Aguzzoli, S.** Collaborative case-based recommendation systems / S. Aguzzoli, P. Avesani, P. Massa // Lecture Notes in Computer Science. 2002. 2416.

ОБ АВТОРЕ

Макеев Григорий Анатольевич, ассист. каф. выч. мат. и киберн. Дипл. инж.-програм. (УГАТУ, 2003). Дис. о системах совместной фильтрации информации.

УДК 621.793

С. Р. ШЕХТМАН

ТЕХНОЛОГИЯ ПОЛУЧЕНИЯ НАНОСТРУКТУРИРОВАННЫХ ЗАЩИТНЫХ ПОКРЫТИЙ

Рассматривается технология создания наноструктурированных покрытий на основе вакуумного ионно-плазменного метода в условиях дополнительной ионной бомбардировки. Приведены исследования свойств покрытий, полученных по предлагаемой технологии. *Вакуумные ионно-плазменные покрытия; ионная бомбардировка; наноструктурированные покрытия; многослойные покрытия*

Применяемые в настоящее время защитные покрытия для целого ряда деталей авиационной техники, работающих в условиях высоких температур, нагрузок и агрессивных сред, не в полной мере отвечают необходимому комплексу требований по их защите. В связи с этим продолжается поиск новых способов и процессов нанесения покрытий в направлении создания композиций, обладающих более высокими эксплуатационными свойствами [1, 2].

В последнее десятилетие карбиды металлов привлекают внимание широкого круга специалистов, занимающихся синтезом этих соединений, изучением их структуры и разнообразных свойств. Высокая температура плавления многих карбидов, их своеобразные механические и физические свойства (большая твердость, абразивная способность, тугоплавкость, пластичность при высоких температурах и др.) обусловливают широкий интерес к покрытиям на их основе. Однако возможности повышения твердости поверхностного слоя при нанесении простых карбидов ограничены [3]. Наиболее перспективными в этом плане являются легированные конденсаты. Карбиды, силициды и карбосилициды металлов обладают уникальным сочетанием высокой твердости, коррозионной стойкости и термодинамической устойчивости, однако получение таких фаз традиционными методами связано с высокой температурой и продолжительностью процесса их синтеза. При этом технологии получения защитных покрытий из таких композиций отсутствуют.

На основе анализа литературных источников [3–6] и предварительных исследований было высказано предположение (гипотеза) о том, что если осуществить в вакууме ионно-плазменное последовательное осаждение веществ системы Ti-C-Si при их одновременной ионной бомбардировке, то возможно получение многослойного покрытия, содержащего такие фазы, как карбиды, силициды и карбосилициды титана, а при последующей термической обработке — регулирование его фазового состава.

Многослойно-композиционные наноструктурированные покрытия достаточно хорошо сопротивляются разрушению в процессе его эксплуатации, благодаря тому, что [4–7]:

- многослойное наноструктурированное покрытие включает в себя чередующиеся тонкие слои переменной твердости, чрезвычайно эффективно тормозит развитие трещины, вследствие создания протяженных полей сжатия (твердые слои чередуются с более мягкими) и барьера на пути ее движения (мягкие тонкие слои);
- снижение модуля упругости покрытия позволяет уменьшить остаточные напряжения в покрытии и соответственно уменьшить градиент напряжений на границе раздела покрытие — основной материал, таким образом, заметно снизив вероятность отслаивания покрытия в процессе эксплуатации;
- многослойная структура покрытия обеспечивает повышенную энергоемкость поверхностных слоев материала основы вследствие ее рассеяния на границах раздела слоев покрытия, что в целом повышает трещиностойкость и вязкость разрушения композиций покрытие — основной материал.

Если толщина покрытия не превышает 100 ... 300 нм, то пленка, как правило, имеет не сплошную, а «островковую» структуру. Многослойное покрытие из слоев, толщина которых не превы-