УДК 004.042 DOI: 10.25045/jpit.v10.i2.12

Шыхалиев Р.Г.
Институт Информационных Технологий НАНА, Баку, Азербайджан
ramiz@science.az

ОБ ОДНОЙ МОДЕЛИ ВИЗУАЛИЗАЦИИ МОНИТОРИНГА КОМПЬЮТЕРНЫХ СЕТЕЙ

Поступила: 26.04.2019 Исправлена: 17.05.2019 Принята: 24.05.2019

Увеличение объема, скорости и гетерогенности состава сетевых трафиков привело к усложнению задач управления современными компьютерными сетями (КС). Для решения этой проблемы требуется новая парадигма управления КС, в основе которой должен лежать непрерывный и эффективный мониторинг. Однако при непрерывном мониторинге КС приходится иметь дело с очень большими объемами данных сетевого трафика. Это в свою очередь приводит к снижению эффективности мониторинга КС, что делает необходимым включение человека в процесс анализа сетевых данных. Для решения этой проблемы в данной статье предлагается модель визуализации мониторинга КС, основанная на методах визуальной аналитики.

Ключевые слова: *мониторинг, сетевой трафик, визуализация, визуальная аналитика, методы визуализации данных, методы добычи данных*

Введение

Сегодня объемы трафика компьютерных сетей, особенно интернет-трафика, безмерно возросли. Это в основном связано с увеличением количества сетевых пользователей, сервисов и приложений. Обеспечение в таких условиях устойчивой работы КС является очень важной и трудной задачей, для решения которой недостаточно только расширения сетевой инфраструктуры. Поэтому появляется необходимость в разработке новой парадигмы управления КС, в основе которой лежит непрерывный и эффективный мониторинг. Однако при непрерывном мониторинге КС приходится непрерывно собирать и хранить большие объемы данных трафика, что затрудняет их анализ.

Несмотря на то, что современные технологии сбора и хранения данных позволяют производить сбор и хранение большого объема сетевых данных, все же возможности их анализа отстают. Поэтому анализ большого объема сетевых данных является очень сложной задачей и делает необходимым участие человека в процессе их анализа.

Для сбора сетевых данных в режиме реального времени в КС широко используются различные системы мониторинга. Вместе с тем ручной детальный анализ результатов мониторинга аналитиками (администраторами КС) остается сложной задачей из-за большого объема данных. В итоге снижается эффективность мониторинга КС, что не позволяет аналитикам использовать эффективные стратегии и принимать обоснованные решения по управлению КС. Выходом из этой ситуации является использование методов визуализации анализа данных, которые позволят аналитикам оперативно включиться в оценку данных. В литературе для визуализации данных были предложены множества различных методов [1–5].

В контексте мониторинга КС визуальный анализ данных позволяет сетевым администраторам через интерактивные визуальные представления лучше понять большие, многомерные и динамически изменяющиеся во времени сетевые данные. Следовательно, сетевые администраторы могут иметь возможность своевременно принимать эффективные решения в критических ситуациях и реагировать на аномальные ситуации в КС. В результате этого могут повыситься скорость и эффективность мониторинга КС.

Существующие сегодня средства мониторинга [6] позволяют визуализировать сетевые данные в самых разных контекстах, однако это осуществляется на основании определенных характеристик сетевого трафика, и при этом не учитывается сложный характер взаимосвязи между данными. Поэтому для эффективного мониторинга КС очень актуально использование методов визуальной аналитики. При этом одной из основных задач является учет сложности и объема данных, а также возможности восприятия человека. Визуализация позволит идентифицировать те проблемы, которые не могут быть идентифицированы в традиционных подходах мониторинга.

Целью статьи является разработка модели визуализации сетевых данных, которая позволила бы осуществлять эффективный мониторинг КС. Для этого предлагается использовать методы добычи данных (англ. data mining) совместно с методами графической визуализации.

Обзор литературы по визуализации мониторинга КС

В области визуализации мониторинга КС предложено множество подходов. Одной из ранних работ в этой области является работа [7], которая была посвящена обнаружению вторжений и злоупотреблений в крупных системах. В этой работе авторами был предложен метод визуализации характеристик взаимодействия хоста сети с другими хостами. Для визуализации используются концентрические кольца, и метод заключается в том, что в центральное кольцо помещается хост, в отношении которого проводится мониторинг. Другие хосты сети, которые инициировали соединения с этим хостом, помещаются в новых концентрических кольцах в соответствии с их расстоянием до исходного хоста в пространстве IP-адресов. При этом эти хосты связываются с исходным хостом с помощью различных типов линий и тем самым кодируются приложения и статус соединения.

На более детальном уровне анализ номеров портов позволяет определить сетевые приложения, которые генерируют трафики. Метод визуализации, использующий номера портов, предложенный в работе [8], основывается на вращающемся кубе, используемом в качестве трехмерной диаграммы рассеяния. При этом к его осям назначались такие атрибуты, как локальное пространство IP-адресов, номера портов и глобальное пространство IP-адресов, и каждый пакет трафика отображался в точку на трехмерной диаграмме рассеяния. Недостатком этого метода является то, что полученные трехмерные диаграммы рассеяния трудно интерпретировать на двухмерной плоскости из-за наложения.

В работе [9] для анализа сетевого трафика сначала осуществляется его визуализация, используя диаграммы рассеяния, диаграммы Ганта или параллельные графики. После этого интерактивно задается шаблон, который необходимо абстрагировать и сохранить с использованием декларативного представления знаний.

Для улучшения управления, мониторинга и безопасности КС с помощью визуализации в работе [10] рассматриваются различные методы. Также анализируются проблемы мониторинга сетей на основе SNMP (Simple Network Management Protocol) и связанные с ним аспекты безопасности. Для улучшения мониторинга сетей предлагается визуализировать в реальном времени процессы маршрутизации, а также потоки сетевого трафика.

В работе [11] для эффективного мониторинга крупномасштабной сети дата центров предлагается архитектура визуализации IP-трафика. Предложенная авторами структура визуализации трафика основывается на алгоритмах машинного обучения, извлечения характеристик IP-трафика и методов графической визуализации данных. Экспериментальные результаты показывают, что предложенный подход позволяет обрабатывать большие потоки данных и предоставляет удобные для пользователя интерактивные графические представления.

В качестве способа мониторинга, анализа и визуализации сетевого трафика могут быть использованы графы рассеивания трафика (Traffic Dispersion Graphs – TDGs) [12]. С помощью TDG авторы моделируют взаимодействия хостов, где ребра графа используются для представления взаимодействия хостов, например, обмен пакетами определенного количества или типа и т.д.

Задача визуального анализа характеристик потоков сетевых трафиков между хостами Интернета является важной и сложной задачей. Поскольку данные потоков сетевых трафиков являются сложными и большими, возникают в реальном времени и содержат сложные взаимосвязи и изменяются во времени. Для решения этой проблемы в работе [13] авторы используют абстрагирование потоков сетевых трафиков на сетевом уровне модели ISO OSI (International Standards Organization Open Systems Interconnection). Визуализация основана на таких атрибутах пакетов, как IP-адреса источника и назначения интернет-хостов и номеров портов. Для визуализации применяется иерархическое радиальное представление данных, которое использует концентрические кольца для отображения распределения сетевых трафиков по атрибутам пакетов.

Как видно из вышеприведенного обзора работ по визуализации мониторинга КС, предложенные в них подходы относятся к решению отдельных конкретных задач мониторинга. В каждом подходе используются определенные сетевые данные, и к задачам мониторинга относятся такие, как мониторинг безопасности сети, мониторинг сетевого трафика, мониторинг взаимодействия хостов КС и т.д. Однако при больших объемах сетевых данных, а также сетей и мониторинга эти подходы могут быть не очень эффективными.

Исходя из вышеперечисленного, можно сказать, что есть необходимость разработки универсальной и гибкой модели визуализации мониторинга КС, которая позволила бы решить различные задачи сетевого мониторинга. Для решения этой проблемы предлагается использовать визуальную аналитику. При этом, в зависимости от задач мониторинга, могут быть использованы те или иные сетевые данные.

Визуальная аналитика

Сегодня для анализа широко используются методы добычи данных, которые позволяют извлекать ценную информацию из сырых данных [14]. Однако при увеличении объема и неоднородности анализируемых данных, а также при необходимости масштабирования алгоритмов в этих методах возникают существенные проблемы. Вместе с тем часто аналитикам становится трудно интуитивно и осмысленно понимать и интерпретировать результаты анализа. В итоге аналитики не могут получать полную информацию, чтобы принять правильные решения. Для решения этой проблемы в литературе были предложены различные методы визуальной аналитики, которые позволяют интегрировать методы автоматического анализа и интерактивной визуализации данных [15].

Визуальная аналитика – это наука об аналитическом мышлении, поддерживаемая интерактивными визуальными интерфейсами [16]. Визуальная аналитика является более широким понятием, чем визуализация данных, и включает в себя визуализацию, анализ данных и человеческий фактор. При этом интерактивная визуализация используется для интеграции знаний и умений аналитиков в процессы анализа данных.

Процесс визуальной аналитики состоит из комбинации автоматического и визуального анализов данных и взаимодействия с человеком для получения знания (рис.1) и делится на несколько этапов [17]. На первом этапе процесса визуальной аналитики производится предварительная обработка данных, типичными задачами которой являются очистка данных, нормализация, группировка или интеграция гетерогенных данных в общую схему. Далее, используя эти данные, аналитик для извлечения знаний может

непосредственно выбрать визуализацию или модели автоматического анализа данных. При выборе аналитиком визуализации (второй этап) производится отображение данных, то есть визуализация данных, полученных на первом этапе. В процессе визуальной аналитики исходных данных знания могут быть получены из этапов визуализации (третий этап), из моделей автоматического анализа, а также из взаимодействия между ними (четвертый этап). При этом, если первоначальная визуализация данных не позволяет получить достаточно информации, то результаты визуализации могут быть повторно использованы для построения модели для автоматического анализа (четвертый этап). А визуализация модели (четвертый этап) может использоваться для проверки результатов этой модели. Таким образом, чередование визуальных и автоматических методов анализа данных позволит аналитикам уточнить и проверить предварительные результаты анализа. Для улучшения результатов анализа включен цикл обратной связи, который позволяет аналитику улучшить выводы в дальнейшем.

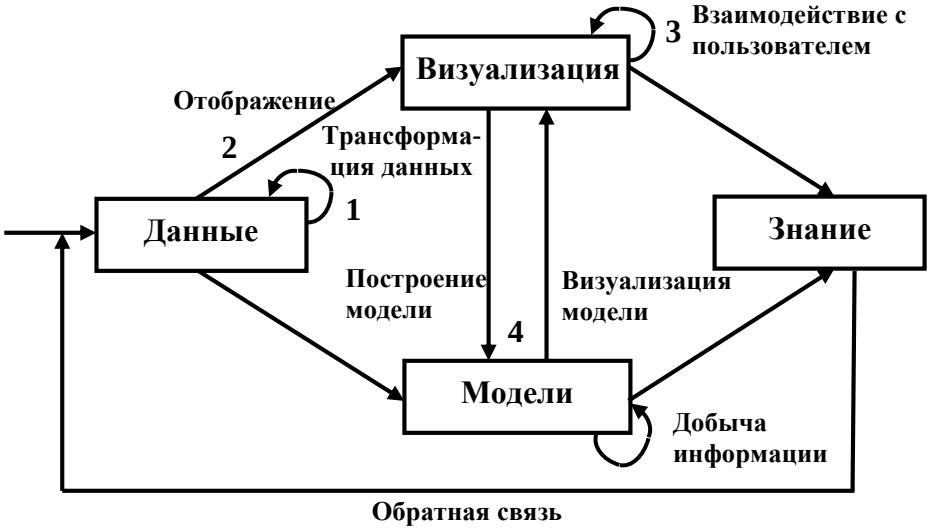

Рис.1. Процесс визуальной аналитики

В общем, визуальная аналитика основывается на множестве смежных научных областей, таких как визуализация, менеджмент данных, добыча и анализ данных, пространственно-временной анализ данных, восприятие и познание человека и методологий оценки [18].

Сегодня популярность визуальной аналитики очень выросла и является активной областью исследований, которая используется в различных сферах. К ним можно отнести такие области, как физика и астрономия, мониторинг климата и погоды, управление в чрезвычайных ситуациях, биология и медицина, бизнес-аналитика, безопасность, анализ сетевого трафика и т.д., где необходимо обрабатывать и анализировать большие объемы данных.

Модель визуализации мониторинга КС

Исходя из предыдущего раздела, для визуализации мониторинга КС предлагается модель, которая является интеграцией методов добычи данных и методов визуализации (рис. 2). Модель состоит из этапов предварительной обработки сетевого трафика, анализа данных, полученных из предварительной обработки трафика, визуализации информации, полученной из анализа данных и обратной связи.

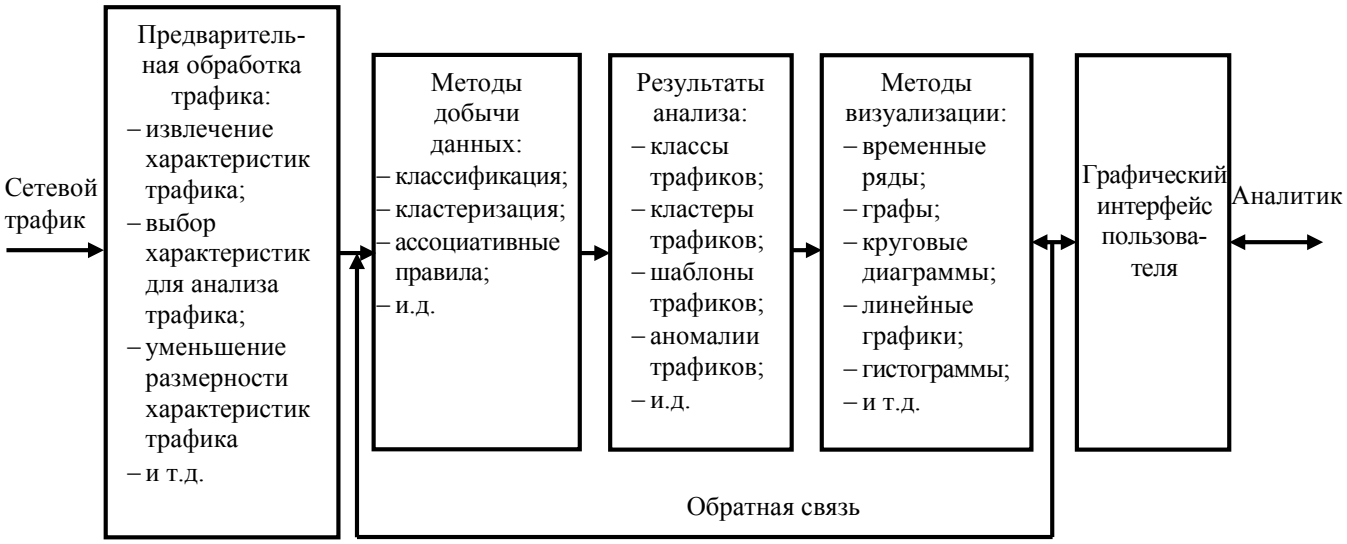

Рис.2. Модель визуализации мониторинга КС

Этап предварительной обработки сетевого трафика предназначен для подготовки данных трафика для дальнейшего анализа. На этом этапе осуществляются извлечение характеристик трафика [19], уменьшение размерности характеристик трафика [20] и выбор характеристик для анализа трафика [21]. При этом выбор наборов характеристик сетевого трафика зависит от целей анализа, то есть от целей мониторинга КС. Например, для классификации сетевого трафика могут быть использованы такие характеристики сетевого трафика, как размер пакетов, IP-адрес источника, порт источника, IP-адрес назначения, порт назначения, тип протокола и т.д. [21, 22].

После предварительной обработки данные сетевого трафика поступают в блок анализа данных для извлечения информации, необходимой для мониторинга, при этом для анализа данных сетевого трафика могут быть использованы различные методы добычи данных и машинного обучения. Для анализа данных сетевого трафика в литературе были предложены множества подходов, основанных на методах извлечения данных [23–26], которые позволят осуществлять мониторинг КС в различных аспектах.

На этапе визуализации результаты анализа данных сетевого трафика, полученных на предыдущем этапе, представляются визуально. Для этого могут быть использованы различные методы визуализации, такие как временные ряды, графы, круговые диаграммы, линейные графики, гистограммы и т.д. При этом выбор тех или иных методов визуализации зависит от видов анализируемых сетевых данных, задач мониторинга, а также от желаемого уровня детализации информации. Визуализация данных сетевого трафика позволит выявить в сетевых трафиках определенные закономерности и тренды. Результаты визуализации представляются аналитику с помощью графического интерфейса пользователя, посредством которого аналитик взаимодействует с системой мониторинга.

Обратная связь предназначена для улучшения результатов анализа на основе визуальной аналитики, что может быть достигнуто изменением параметров алгоритмов, и аналитики могут наблюдать результаты этих изменений визуально.

Заключение

В данной статье предлагается модель визуализации мониторинга КС. Предложенная модель основывается на интеграции методов добычи данных и графической визуализации данных. При этом методы добычи данных используются для извлечения из сетевого трафика информации, необходимой для мониторинга КС, а с помощью графического

представления эта информация визуализируется. При этом выбор методов извлечения данных и методов визуализации зависит от целей мониторинга КС.

Предложенная модель позволит осуществлять аналитику сетевого трафика и визуализацию мониторинга КС в различных аспектах.

Литература

1. Khan M., Khan S.S. Data and Information Visualization Methods, and Interactive Mechanisms: A Survey // International Journal of Computer Applications, 2011, vol.34, no.1, pp.1–14.
2. Dzemyda G., Kurasova O., Zilinskas J. Multidimensional data visualization. Methods and applications, 2015, 252 p.
3. Wang L. Big Data and IT Network Data Visualization // International Journal of Mathematical Engineering and Management Sciences, 2018, vol.3, no.1, pp.9–16.
4. Lengler R. and Eppler M.J. Towards a periodic table of visualization methods for management / Proceedings of the IASTED International Conference on Graphics and Visualization in Engineering, 2007, pp.83–88.
5. Becker R.A., Eick S.G., Wilks A.R. Visualizing Network Data // IEEE Transactions on Visualization and Computer Graphics, 1995, vol.1, no.1, pp.16–21.
6. Skurek J.A Survey of Tools for Monitoring and Visualization of Network Traffic, Bachelor's Thesis, Masaryk University, 2015.
7. Erbacher R.F., Walker K.L., and Frincke D.A. Intrusion and misuse detection in large-scale systems // IEEE Computer Graphics and Applications, 2002, vol.22, no.1, pp.38–48.
8. Lau S. The spinning cube of potential doom // Communications of the ACM, vol. 47, no.6, 2004.
9. Xiao L., Gerth J., and Hanrahan P. Enhancing visual analysis of network traffic using a knowledge representation / Visual Analytics Science and Technology (VAST), 2006, pp. 107–114.
10. Miller K.B. and Brandon E.R., Improving Network Monitoring and Security via Visualization, 2016, https://arxiv.org/pdf/1511.08795
11. Elbaham M., Nguyen K.K., Cherie M. A Traffic Visualization Framework for Monitoring Large-scale Inter- DataCenter Network / 12th International Conference on Network and Service Management, 2016, pp.277–281.
12. Iliofotou M., Pappu P., Faloutsos M. Network Monitoring using Traffic Dispersion Graphs (TDGs) / Proceedings of the 7th ACM SIGCOMM conference on Internet measurement 2007, pp. 315–320.
13. Keim D.A., Mansmann F., Schneidewind J., Schreck T. Monitoring Network Traffic with Radial Traffic Analyzer / IEEE Symposium On Visual Analytics Science And Technology, 2006, pp.123–128,.
14. Han J. and Kamber M., Data mining: concepts and techniques. Morgan Kaufmann, 2006.
15. Sun G.D., Wu Y.C., Liang R.H. et al. A survey of visual analytics techniques and applications: State-of-the-art research and future challenges // Journal of computer science and technology, 2013, vol.28, no.5, pp.852–867.
16. Thomas J. and Cook K. Illuminating the Path: Research and Development Agenda for Visual Analytics, IEEE-Press, 2005, 184 p.
17. Keim D.A., Mansmann F., Stoffel A., Ziegler H. Visual Analytics, Springer, 2009. Encyclopedia of Database Systems.
18. Keim D.A., Andrienko G., Fekete J.D., Gorg C., Kohlhammer J., and Melancon G. Visual Analytics: Definition, Process, and Challenges / Information Visualization, LNCS 4950, pp. 154–175, 2008.

19. Шыхалиев Р.Г. // О методе извлечения классификационных признаков сетевых трафиков на основе анализа сигналов // Проблемы Информационных Технологий, 2019, №1, с.78–86.
20. Шыхалиев Р.Г. Об одном методе сокращения размерности анализируемых признаков сетевых трафиков, используемых для мониторинга компьютерных сетей // Телекоммуникации, 2011, №6, с. 44–48.
21. Шыхалиев Р.Г. Анализ и классификация сетевого трафика компьютерных сетей // Проблемы Информационных Технологий, 2010, №2, с.15–23.
22. Шыхалиев Р.Г. Об одном методе классификации трафика компьютерных сетей // Проблемы Информационных Технологий, 2014, №2, с.59–67.
23. Adibi S. Traffic Classification – Packet-, Flow-, and Application-based Approaches // International Journal of Advanced Computer Science and Applications, 2010, vol.1, no.1, pp.6–15.
24. Lee I. W., Fapojuwo A.O. Data Mining Network Traffic / Canadian Conference on Electrical and Computer Engineering, 2006, pp.148–152.
25. Prangchumpol D. A. Network Traffic Prediction Algorithm Based On Data Mining Technique // World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering, 2013, vol.7, no.7, pp.999–1002.
26. Joshi M.R., Hadi T.H. A Review of Network Traffic Analysis and Prediction Techniques, 2015, https://arxiv.org/abs/1507.05722

UOT 004.042

Şıxəliyev Ramiz H.
AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
ramiz@science.az

Kompüter şəbəkələrinin monitorinqinin vizuallaşdırılmasının bir modeli haqqında

Şəbəkə trafiknin həcminin, sürətinin və heterogenliyinin artması müasir kompüter şəbəkələrinin (KŞ) idarəetmə məsələlərinin çətinləşməsinə səbəb olmuşdur. Bu problemi həll etmək üçün fasiləsiz, effektiv monitorinqə əsaslanan yeni bir KŞ idarəetmə paradiqması tələb olunur. Bununla belə, KŞ-nin fasiləsiz monitorinqi zamanı çox böyük həcimdə şəbəkə trafiki məlumatları toplanır. Bu, öz növbəsində KŞ-nin monitorinqinin səmərəliliyinin azalmasına gətirib çıxarır ki, bu da insanın şəbəkə məlumatlarının təhlili prosesinə qoşulmasını zəruri edir. Bunun üçün, bu məqalədə KŞ-nin monitorinqinin vizallaşdırılması üçün vizual analitika metodlarına əsaslanan model təklif edilir.

***Açar sözlər**: monitorinq, şəbəkə trafiki, vizuallaşdırma, vizual analitika, verilənlərin vizuallaşdırılması metodları, verilənlərin intellektual analizi metodları*

Ramiz H. Shikhaliyev
Institute of Information Technology of ANAS, Baku, Azerbaijan
ramiz@science.az

One model of visualization monitoring of computer networks

The increase in the volume, speed and heterogeneity of the network traffic composition has led to the complication of the management tasks of modern computer networks (CN). To solve this problem, a new CN management paradigm is needed, which should be based on continuous, effective monitoring. However, with continuous monitoring of the CN, we have to deal with very large amounts of network traffic data. This, in turn, leads to a decrease in the efficiency of monitoring the CN, which makes it necessary to include a person in the process of analyzing network data. For this, in this article proposes a model visualization of the CN monitoring based on the methods of visual analytics.

***Keywords**: monitoring, network traffic, visualization, visual analytics, data visualization methods, data mining methods.*