BLOG@CACM

The *Communications* website, http://cacm.acm.org, features more than a dozen bloggers in the BLOG@CACM community. In each issue of *Communications*, we'll publish selected posts or excerpts.

twitter
Follow us on Twitter at http://twitter.com/blogCACM

DOI:10.1145/3530685 http://cacm.acm.org/blogs/blog-cacm

The Role of Math in IT Education

Andrei Sukhov considers why and how the foundations of teaching mathematics for information technology specialties need to be revised.

Andrei Sukhov

How to Teach Mathematical Disciplines for IT Specialties

January 19, 2022

https://bit.ly/3HMpRoX

The role of mathematical disciplines in IT education is difficult to overestimate.

First of all, mathematics helps to develop algorithmic thinking, as mathematical theories are based on abstract concepts. Mastering any section of fundamental mathematics allows you to operate with abstract concepts to find new patterns.

However, undergraduate IT students often doubt the need to study mathematical disciplines. Their main argument is that the study of mathematics takes a lot of time and effort, but gives very little for the practical application of the acquired knowledge. Therefore, the time spent on studying mathematics can be better used for special courses in IT disciplines.

That is why I decided to discuss the problems of teaching mathematical disciplines, and to propose a number of new approaches.

Note that knowledge of mathematics is particularly in demand at the graduate level. My foreign colleagues often ask me to recommend a programming student to graduate school, but with a mandatory knowledge of mathematics. In Russia, computer science faculties have several areas of study, such as applied mathematics and physics. The curriculum for these specialties contains a considerable portion of programming, and various areas of mathematics and physics. However, the popularity of such areas of training is declining, as is the average level of incoming applicants.

Thus, a revision of the foundations of teaching mathematics for IT specialties is required.

First, we need to have a small number of required courses. Their composition should be discussed, but in my opinion, it is necessary to have only three basic courses. These are differential and integral calculus, basics of algebra, and probability theory with mathematical statistics.

Most of the mathematical theories should be taught as part of new comprehensive courses. Such courses can be devoted to a narrow area of IT technologies and should contain the following three main components:

- basic information from the section on fundamental mathematics;
- formulation of an applied problem and its solution using the studied mathematical approaches, and
- practical implementation of the obtained solutions by means of IT technologies.

The main thing such courses should teach is to go through all the stages from abstract fundamental knowledge to a full-fledged application. Moreover, many applications, particularly the best ones, are based on fundamental knowledge. However, students often have no idea how to get through this path from start to finish. They also do not understand what knowledge is required. At best, they are taught applied mathematics and its application to solve some problems.

In principle, many sections of fundamental mathematics are of applied

importance; however, it is quite difficult to develop a full-fledged comprehensive course. This difficulty can be attributed to the large number of consistent conclusions that need to be presented in such a course. Such a presentation requires a broad outlook from the author.

At present, I have begun preparing such a course, using which I will illustrate the main features of this approach. Let us now dwell on the main provisions and content of the course "Traffic model of the backbone network and justification of the threshold method for detecting DDoS (distributed denial-of-service) attacks."

Such a course is fully consistent with the idea of an integrated approach, and contains the following main components:

- queuing theory,
- its applications for describing traffic on the backbone network,
- applied traffic model,
- determination of the abnormal state of the network, and
- substantiation of the threshold value method for identifying sources of DDoS attacks.

Queuing theory is a branch of probability theory based on the problem of the processes of death and reproduction. Advanced research in this area has led to several applications for a wide range of real socioeconomic and demographic processes. However, the first field of application of queuing theory was telecommunications. The main provisions and methods of analysis of the queuing theory in telecommunications are brilliantly presented in textbooks.[1,2] These textbooks are used as a methodological basis for the first component of the course, and should be recommended to students as additional literature.

The next step in building a traffic model on a section of the backbone network was made in the article.[3] By means of the queuing theory, expressions were found for the average traffic on a backbone link and for its variations at short timescales. Note that the generalizations were made at the level of flows, not packets. In particular, the expressions for traffic include the average flow size and the average rate of appearance of new flows on the studied backbone section.

"In my opinion, it is necessary to have only three basic courses: differential and integral calculus, basics of algebra, and probability theory with mathematical statistics."

However, obtaining expressions for traffic and its short-term variation does not yet mean building a full-fledged traffic model. Such a model should define the area of normal operation, as well as highlighting anomalous network states. All these goals can be achieved if the network state is described by two variables: the number of active flows in the network section and the link load in bits per second.[4] Then, the set of network states will be represented by a set of points on the plane, with the abscissa equal to the number of active flows and the ordinate representing the link load.

On this plane, we can build a curve from the averaged values and select a straight part on it corresponding to the operating mode. Depending on the quantile, it is possible to define a parabolic region with a central axis in the form of this straight line. Points corresponding to network states will fall into this area. If several successive states go beyond this area, then we can talk about the anomalous state of the network.

The conducted experiments have shown that the values for some network variables increase many times during DDoS attacks. This fact was first discovered for the number of active flows generated by a single external IP address. Subsequently, it turned out that such variables include incoming TCP and UDP traffic and the number of calls to the Web or proxy server.

In Sukhov, Andrei M., et al.[5], it was proven that for all of these variables, it is possible to find a threshold value. If the threshold is exceeded, we should talk about a DDoS attack. This article describes how to find threshold values and formulates rules for determining the attacking IP addresses. Thus, the theoretical part of the course can be considered to be complete.

In the practical part of the course, it is necessary to create tools to determine the beginning of an attack, such as the IP addresses from which an attack is being carried out, and develop ways to block attacking traffic. The practical implementation of these tools can be carried out on the basis of a number of technologies. These can be Linux utilities, SDN modules, NetFlow, and sFlow collectors.

Students must independently choose a technology for the practical implementation of a theoretical model and justify their choice. It is also necessary to propose a method for restricting traffic from attacking IP addresses, as the theoretical model does not answer this question. The developed software should detect an attack as quickly as possible and start blocking the attacking traffic, and after the attack stops, remove all the restrictions.

I hope to have this course ready by fall 2022 and to offer it as an elective course. In principle, I would very much like to hear comments on the proposed course, as well as on the concept of comprehensive courses in the study of mathematics. In addition, I would like to offer cooperation in the development of topics for comprehensive courses for everyone. Please respond in the comments if you have any suggestions or comments.

Footnotes

1. Kleinrock, L. *Theory, volume 1, Queueing Systems*. (1975).
2. Gnedenko, B.V. and Kovalenko, I.N. *Introduction to Queueing Theory*. Birkhauser Boston Inc., 1989.
3. Barakat, C. et al. Modeling Internet backbone traffic at the flow level. *IEEE Transactions on Signal Processing* 51.8 (2003), 2111–2124.
4. Sukhov, A.M. et al. Active flows in diagnostic of troubleshooting on backbone links. *Journal of High Speed Networks* 18.1 (2011), 69–81.
5. Sukhov, A.M., Sagatov, E.S., and Baskakov, A.V. Rank distribution for determining the threshold values of network variables and the analysis of DDoS attacks. *Procedia engineering* 201 (2017), 417–427.

Andrei Sukhov (asukhov@acm.org) is a professor of HSE University and a senior member of ACM.