Cooperative AntiVirus

New Mexico
Supercomputing Challenge
Final Report
April 9th, 2011

Team 19
CEPI #1

Team Members:
Michael Szanto
Devin Hayes

Teacher:
Jerry Esquivel

Summary

For our project, we wanted to prove the positive effects of a cooperative virus signature database. We attempted to prove this by creating a computational model using real life inputs of virus detection data gathered from popular AntiVirus programs. The first step to achieving our goal was to gather data on detection rates of popular AntiVirus programs. To achieve this, we used 3 different computers, each with 3 different operating systems. One computer with Windows XP, Windows Vista, and Windows 7.

To gather our sample viruses we downloaded hundreds of suspicious files via P2P networks, and we tested for self replication and unauthorized outbound connections. Once we identified a few suspected viruses, it was time to start testing!

On each of the test computers, we installed a baseline security program that restores the system to a base image upon restart, deleting all changes made while the computer was on. We then installed one AntiVirus software program, and introduced a virus to see if it was detected. We recorded our data, restarted the computer, and installed another test AntiVirus. Thanks to our baseline security program, by restarting between each test we were able to effectively able to eliminate unnecessary complications such as successful infection by a virus. Using the data we obtained above, we created a simulation that can model the spread over the internet of most types of viruses in existence today.

The Problem

Everyday thousands of computers are unnecessarily infected by viruses that are already well known by one or more AntiVirus providers. This is because AntiVirus providers don't share when they detect a new virus, they keep it to themselves because they all strive to be the best. For example, Mcafee could know about a virus that Norton doesn't and the customers of Norton would all be vulnerable.

The Solution

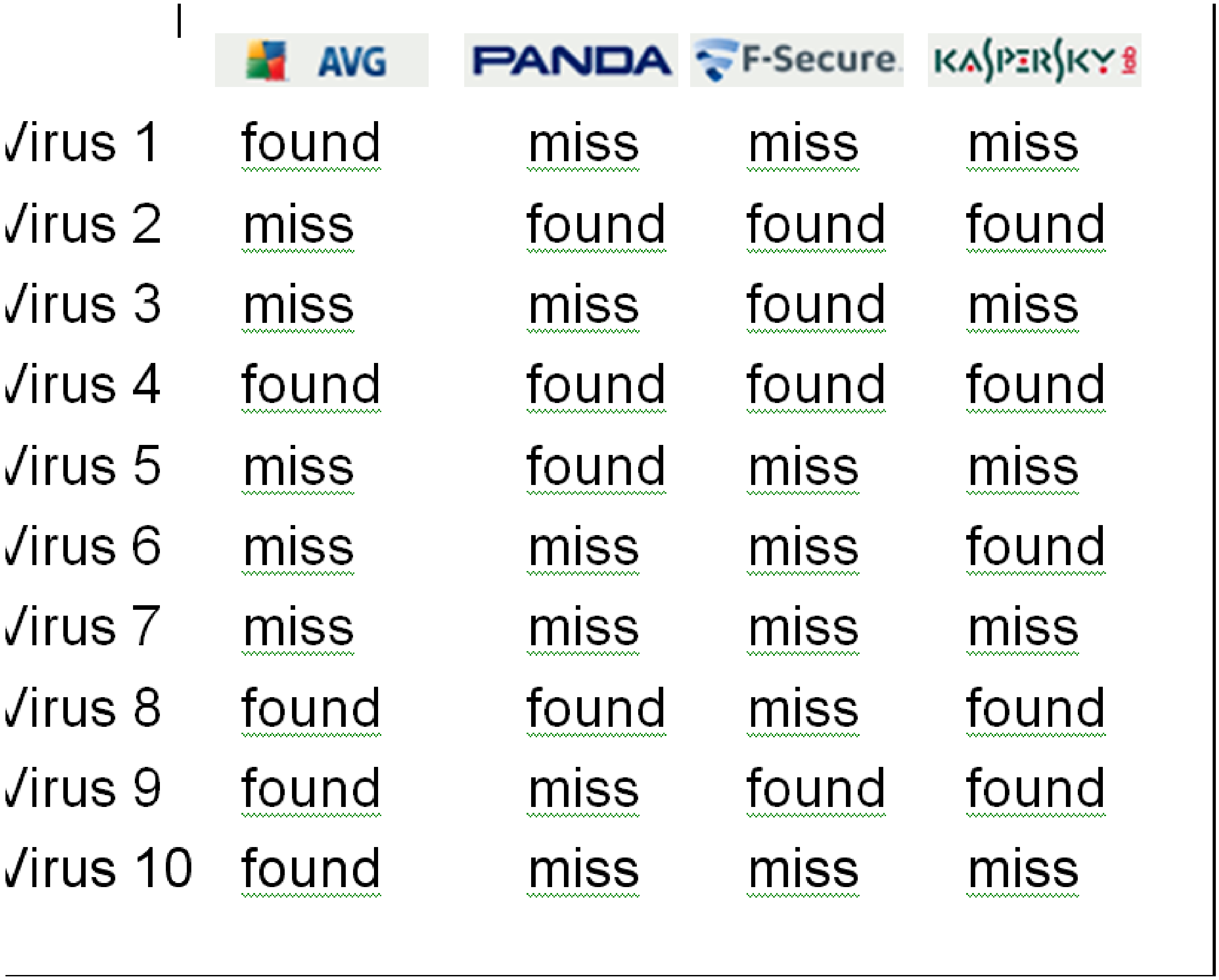
The solution would be for AntiVirus software providers to share file signature detections and sample files. This would allow for more immediate detection of viruses and would benefit all end users.

Program Code

At this time we do not have 100% functional code, but our program does process 90% of our data inputs with surprising results.

The Results

Viruses Detected by Test AntiVirus Software:



	AVG	PANDA	F-Secure	KASPERSKY
/irus 1	found	miss	miss	miss
/irus 2	miss	found	found	found
/irus 3	miss	miss	found	miss
/irus 4	found	found	found	found
/irus 5	miss	found	miss	miss
/irus 6	miss	miss	miss	found
/irus 7	miss	miss	miss	miss
/irus 8	found	found	miss	found
/irus 9	found	miss	found	found
/irus 10	found	miss	miss	miss

Conclusion

In conclusion, our hypothesis was correct. After testing nearly all permutations of theoretical viruses our simulation was programmed to simulate, all cases that involved a shared database always provided a varying margin of better protection.