UNIVERSITY OF CALIFORNIA
RIVERSIDE

Fighting Spam, Phishing and Email Fraud

A Thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science

in

Computer Science

by

Shalendra Chhabra

December 2005

Thesis Committee:

Professor Dimitrios Gunopulos, Chairperson
Professor Vana Kalogeraki
Professor Eamonn Keogh
Professor Mart Molle
Dr. William S. Yerazunis

Copyright by
Shalendra Chhabra
2005

The Thesis of Shalendra Chhabra is approved:

_____

_____

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

Dreaming big with a drive to excel requires immense motivation and a positive attitude. It takes a lot of courage to follow one's heart and chase one's dreams. Great things are never done without hard work and sacrifice.

I am grateful to my *parents* for teaching me to dream big, for blessing me with the courage to follow my heart, for inspiring me to work hard, and for instilling the spirit of sportsmanship in me. They brought me up in a very optimistic environment, giving me a remarkable ability to judge good from bad. I am grateful to my *sisters* for being the best *sisters* in the world. Every success in life would not have been possible without their sacrifice, help, and encouragement. They taught me that *winners don't do different things but they do things differently*. I am grateful to my *grandparents*, my *uncle*, my *aunt*, my *cousins*, and other members of my *family* for their love, affection, and support during difficult times.

I am grateful to all my *friends, teachers*, and *roommates* for enriching my life.

I am highly indebted to members of the *Graduate Division* at the *University of California, Riverside (UCR)*, for providing me with an opportunity to pursue graduate studies at this temple of education.

This thesis is a compiled version of the cumulative knowledge and wisdom gained throughout my two years at the *Department of Computer Science and Engineering* at *UCR*; internships at *Mitsubishi Electric Research Laboratories (MERL)*; presentations at *MIT, Stanford University,* and *Cisco Systems*; interactions with people from around the world; and fruitful discussions from my interviews at *Microsoft, Cisco Systems, Intel*, and *Google*.

Many people have contributed to this thesis directly or indirectly. It's an honor to have come across so many wonderful people.

I am highly indebted to my advisor, *Professor Dimitrios Gunopulos*, for motivating, guiding, and helping me channel my creative energy at times when I was unable to focus. This thesis would not have been possible without his encouragement, wisdom, feedback, and support. His counselling kept me pointed in the right direction, and his periodic feedback protected me from pitfalls.

I am highly grateful to members of my thesis committee—*Professor Vana Kalogeraki* and *Professor Eamonn Keogh* for providing guidance, encouragement, and valuable insight. Communication with them has always brought cheerful spirits and inspiration in me.

I am highly indebted to *Dr. William "Bill" S. Yerazunis (Captain Crash of Junkyard Wars)* from *MERL* for being an amazing mentor and an incredible collaborator during 2004 and 2005. He is the smartest person I have ever come across. Working with him has added tremendous value to my life. Words are not enough to thank *Bill* for the role he has played in my successive successes.

I am highly indebted to *Professor Mart Molle* for his continued, unparalleled support as my graduate advisor. He is a very caring, understanding, excellent teacher who imparts novel ideas to his students. It was during one of his class projects in *Advance Computer Networks* class during Fall of 2003 that led me to the arena of spam-fighting. Words are not enough to express my gratitude for him.

I am grateful to *Dean Dallas L. Rabenstein* for approving continued support for my Master's Degree at *UCR*.

I am thankful to *Professor Stefano Lonardi* for cheering me with his wonderful smile. Fruitful discussions, and communication with him add more confidence in my life.

I am thankful to *Professor Harry C. Hsieh* for being my academic advisor in the early days at *UCR*; to *Professors Chinya V. Ravishankar, Satish Tripathi, Michalis Faloutsos, Christian R. Shelton*, and *Jun Yang* for very enlightening discussions, and to *Department Chair Professor Thomas Payne* for providing valuable teaching tips and resources. Many thanks to members of the Data Mining Lab—*Sharmila Subramaniam, Mirella Moro, Zografoula Vagena, Song Lin, Petko Bakalov, Marios Hadjieleftheriou, Dimitris Papadopoulos, Demetrios Z. Yazti*, and *Christina Charalambidou* for their help and a good time. Many thanks to the department's administrative and management staff—*Terri Phonharath, Kathy Vu, Lilian Kang, Mimi Trinh, Tiffany Arakai, Nicole Smidt, Dulce Shipley, Janice Leslie*, and *Emily Papavero*. Many thanks to *Kelly Hinosowa, Harrison Leong* and *Tatiana Rakic* for being the best international advisors. Many thanks to members of the systems group—*Victor Hill, Benjamin Arai, Conley Read, Craig Boucher, Jacob Lewallen, Benjamin Arai*, and *Chuck Boucher*. Many thanks to *Brian Linard, John Cortes, Ronald Feliciano,* and my *students* from *Computer Organization and Assembly Language* class for a good time.

I am highly grateful to *Dr. Bhiksha Raj* from *MERL* for being an incredible mentor and a great friend. His guidance and wisdom brought a smile to my face every day during the summers of 2004 and 2005. His inspirational notes crystallized my conviction to excel in constraints. This thesis would not have been in its present form without his constant support, motivation, and guidance.

I am grateful to *Paul Graham* for playing a major role in my success, for inspirational communication, for writing wonderful essays, for organizing the *MIT Spam Conference*, and for hosting dinners for anti-spam experts every year. It was because of the MIT Spam Conference of 2004 that I eventually met *Bill Yerazunis*.

I am thankful to *Joe Marks, Bent Schmidt-Nielsen,* and others from *MERL*; to *Rita Singh* from *Haikya Corporation* for great advice, and to *Christian Siefkes (Freie Universität, Berlin)* and *Fidelis Assis (Embratel)* for being great collaborators.

I am grateful to *Joshua Goodman* for introducing me to the *MSN Safety Team,* for providing me with an opportunity to volunteer at the *Second Conference on Email and Anti-Spam (CEAS 2005),* and for invaluable feedback and tips. I am grateful to *Manav Mishra* for stimulating discussions and feedback at the *Email Authentication Implementation Summit 2005, New York City,* and to *Robert Rounthwaite* for intellectual conversations during lunch at *CEAS 2005*. I am grateful to *Kumar Chellapilla* for invaluable feedback and tips. Many thanks to *Harry S. Katz, Kris Iverson, John Scarrow, Ning Zhang, Geoff Hulten, Collene Georgia, Jeff Aylesworth,* and others from *Microsoft Corporation* for brainstorming conversations, and for their time and help during relocation.

I am grateful to *Jim Fenton* for introducing me to the *Network and Spam Solutions Team* at *Cisco Systems* and for great discussions. Many thanks to *Shamim Pirzada, Sanjay Pol, Michael Thomas,* and others from *Cisco Systems* for their time and fruitful conversations.

Many thanks to *Michael Ripley, Gary Graunke, Michael Andre, Don Whiteside, Gary Mittelstaedt, Keith Shippy,* and *Janie C. Mason* from *Intel* and to *Charles Haynes, Kristin Kassaei,* and members of the Gmail team from *Google* for their time and very fruitful face-to-face discussions.

Many thanks to *lawyers Matthew Prince (Unspam Technologies), Jon Praed (Internet Law Group),* and *Aaron Kornblum (Microsoft)* for motivation, guidance, transcripts and relevant resources for legal actions against spam.

Many thanks to *Jonathan Zdziarski (DSPAM)* and *Richard Jowsey (Death2Spam)* for reviewing draft versions of this thesis. Many thanks to *Eric S. Johansson (CAMRAM)* for providing information and feedback on CAMRAM, and to *Matt Blumberg* and *Andy Sautins (Return Path)* for feedback on the *Bonded Sender* section. Many thanks to *Erik Brown* for designing the logo for the thesis. Many thanks to *Shawn R. Lesniak* and *Amir Friedländer* for making proofreading super-fun.

I am thankful to *Laird A. Breyer (University of Lancaster, UK); Regu Radhakrishnan, Ronald Johnson, Ajay Divakaran (MERL); David Silver (MarkMonitor); Katherine Bretz, Des Cahill, Doug Warren (Habeas); Ether Dyson (CNET); Miles Libbey (Yahoo); P. Oscar Boykin (University of Floria); John Graham-Cumming (Electric Cloud); Larry McGrath, Vickie Park, Dallas Johnson, Mitch Boretz, Monica Wicker, Kara Oswood, Karen Smith,*

Abraham Artuz (UCR); Larry Reeker (NIST); Jonathan Oliver, Chad West, Leon Hilton (MailFrontier); Daniel Quinlan (IronPort Systems); Gordon Cormack (University of Waterloo); Mike Linksvayer (Creative Commons); and *Rebecca Wetzel* for very fruitful discussions.

I am grateful to my previous mentors, members of the security group at the *University of Milan—- Professor Pierangela Samarati, Sabrina De Capitani, Stefano Paraboschi, Ernesto Damiani*; *Professor Jean Goubault Larrecq* at the *Laboratoire Spécification et Vérification, Ecole Normale Superieure De Cachan, France*; *Dr. N. Raja* at the *Tata Institute of Fundamental Research, India*, and *Professor S. C. Gupta* at the *Institute of Technology, Banaras Hindu University, India*, for exposing me to research during my undergraduate studies.

I am thankful to all my *friends* for standing by me during the worst times. To express my gratitude for my *friends* would require far more pages than those occupied by the contents of this thesis. My movie, *The BackBenchers*, may prove to be a better medium for this task, provided it is ever finished.

I am thankful to my former roommate *Casey Christine Hoover (UCR)* for being the best roommate in the world. Without the encouragement of this precious person, I would never have discussed and conducted anti-spam measures with the *Computing and Communications (C&C)* group at *UCR*. Her good words—**"uniqueness is a virtue, lead by example, actions speak louder than words, good leaders know how to ask for help, and you will be fine"**—created an optimistic air in the house and pumped me with confidence for presentations at the MIT Spam Conference (2005), for challenging interviews at big corporations, and during tough times, especially during my car accident.

Finally I am thankful to God for showering His love and affection, for bringing wonderful people into my life, and for everything I have achieved to this day.

With such great blessings and positive energy, a revolution to put an end to the spammers' business model is in the cards. It begins on the next page...

CRM114 - the Controllable Regex Mutilator
CRM114 - the Controllable Regex Mutilator
Fighting Spam, Phishing and Email Fraud
Revolution is in the Cards
SPAM
PHISHING
EMAIL FRAUD