Resisting Whitewashing: A Comparative Study of Fixed Identity, Pseudonym and Social Identity

Yijia Xu

School of Computer Science
McGill University
Montreal, Canada

October 2011

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science.

2011/10/26

Abstract

The Internet has broadened the meaning of identity. In real life, we normally have fixed identities. Earlier on the Internet, pseudonym was invented to identify users. As social networks came into play, social identity was introduced. While fixed identity is the most secure one, it restricts the freedom people have on the Internet. Although pseudonym is the easiest one to create among the three identity types, it makes it equally easy for malicious users to cheat on other people. One way malicious users can perform attacks on others with pseudonym is by whitewashing, where the malicious user takes advantage of the victim, with either pre-designed plots or simply by breaking pre-defined rules, and disappear on the network but later re-join the network with a new identity so that no one would know about his previous activities. Social identity differs from the other two by making use of the relations between two people and protecting both the individual user's privacy and the organization's security.

This thesis studies the effect of whitewashing under different identity types and compares their behaviors of resisting whitewashing. We use game theory to model the process of whitewashing and compute its effect upon the whole population for each identity type. In most of the cases, social identity is better at eliminating whitewashers. Besides, the Matlab simulation experiments reveal additional interesting facts about the three identity types that might shed light on future identity management schemes.

Résumé(Français)

Le sens du mot “identité” s’élargit dans le contexte d’Internet. Dans la vraie vie, sauf dans des cas très particuliers, nous avons une identité fixe. Très tôt sur Internet, on a eu recours aux pseudonymes pour identifier les usagers. Avec l’évolution des réseaux sociaux est venue l’identité sociale (“social identity” en Anglais). Bien qu’elle représente le mode d’identité le plus sécuritaire, l’identité fixe impose des limites importantes qui la rendent difficile d’utilisation sur Internet. Parmi les trois modes d’identité, le pseudonyme est le plus facile à créer, mais il permet aussi aux usagers malveillants de facilement abuser de la confiance des autres. Suite aux actes malhonnêtes (arnaques ou non respect des règles établies, par exemple), un usager malveillant peut changer de pseudonyme pour blanchir son identité (“whitewashing” en Anglais). Il disparaît donc du réseau pendant un certain temps pour réapparaitre plus tard sous une nouvelle identité aucunement liée à l’ancienne. L’identité sociale diffère de l’identité fixe et des pseudonymes en se basant sur les relations entre deux personnes tout en protégeant la vie privée de l’utilisateur et la sécurité de l’organisation.

Cette thèse analyse les effets du blanchissage selon le mode d’identité utilisé et compare les différents comportements de résistance au blanchissage. La théorie des jeux est utilisée pour modéliser le blanchissage et calculer son effet sur la population pour chaque mode d’identité utilisé. Dans la majorité des cas, l’identité sociale est plus apte à éliminer les blanchisseurs. Des simulations Matlab révèlent aussi des faits intéressants au sujet des trois modes d’identité qui pourraient aider à l’élaboration de futurs systèmes de gestion d’identité.

Acknowledgments

First and foremost, I need to express my gratitude to Prof. Muthucumaru Maheswaran for his guidance on my research. His wide span of knowledge and logical thinking always inspired me during my research and thesis writing. Without his supervision, it could have been harder to complete my thesis.

I owe my thanks to everyone in the Advanced Network Research Lab for their interesting opinions. My warm thanks to Bader Ali, Amin Ranjbar and Arash Nourien for their friendly help. I'd like to acknowledge Kang Wang, Ning Jia, Fugui Tang, Varun Maheshwari and Yiwei Shi as well, who generously shared interesting discussions on my topic and created a pleasant working environment.

Everyone I met during my M.S.c study, David Kawrykow, Tristan Ratchford, Annie Ying, Theresa Deering, Xiaoxi Dong, Yancheng Xiao, Wenyang Ku, Xiaohu Xie also has my sincere thanks. Former lab mates, Xi Chen, Hanqiang Cheng, Yuan Jin, Yi Zhang and Jianxun Dang together with former classmate who became my graduate friend at McGill now, Yue Gao helped me in every way maintaining a wonderful life through these two years.

I also feel indebt to my mother Youmei Wang, my father Sen Xu for their unfailing support and encouragement all the time, without them my graduate study would not be possible. And my family members within two "hops", who are always so nice to talk to, should have my special thanks. Besides, I appreciate the support from my friends, Xi Lu, Xiaoyin Wang, Yuwei Wang, Yu Feng, Tianshuang Liu, Ruolan Xu, Peilu Xin, Zheng Zhou, Yang Song, Xin Liu and Hongchi Zhang.

Contents

List of Figures

List of Tables

Chapter 1

Identity

Identity can be considered as a certificate or a license that one must have in order to be included in an organization or to be entitled to certain rights. It usually consists of a set of personal information to distinguish the owner from others. The most common identity we have in real life is the ID card we get as a citizen of a country. The purpose of this identity card is to claim that one belongs to a country and is a legal resident; hence, he has all sorts of rights by law. But the concept of identity dates back to ancient times, long before constitutional law was invented. To explain how identity came into being involves some knowledge of philosophy. The starting point of identity was when human first differentiated among themselves and started recognizing "self". But identity has become more than just a tool to distinguish people. Akerlof and Kranton include identity into economic models of behavior and find that identity can affect individual interactions [1]. Eaton, Eswaran and Oxoby argue that we need identity or rather an identity system for the sake of survival [2].

To create a national identity, a name must be registered followed by one's birth date and a lot of other information. In some countries, even finger prints are collected to issue a passport. One might think that his name, birth date, home address and which company he's working for and other information are of no use to any other people and won't bring any benefit to a thief. But there are many chances a thief can gain profit from our information. For instance, a lot of people have received email spam and phone spam. According to the Message Anti-Abuse Working Group, in 2010, email spam took up 90% of total emails. The annoying spam is just one of the reasons why we need to protect our personal information. Sometimes, you reveal your contact information or working information to a party that you trust, like a bank or a bookstore. A staff working at the bank or bookstore will collect a lot of customers information and he or she can sell it to another company that wants to promote their products. This is only the simplest way that your information can be of benefit to someone else.

1.1 Identity theft and fraud

Identity theft or impersonation refers to the act that someone claims to be someone else and cheats on other people. For instance, a thief can copy one of your close friends' profile on a social network and ask for money from you. Or someone claims himself as a representative from a corporation or a well-known company and has an inside information source that can bring possible profit that is double the amount of profit on the market. If you go and check this representative does exist in real life. Then he can ask you for initial investment and actually he might receive payments from a lot of people. He'll disappear after he got enough money. In this case, not only ordinary people are victims of identity theft, organizations can suffer from identity theft as well. In a sense, it's a lot like spoofing attacks on computer networks. As this kind of identity theft has been reported a lot of times, people are becoming more aware of it. The most popular illustration of identity theft might be the movie Catch Me If You Can, which is a movie based on the true story of Frank Abagnale Jr. who assumed more than eight identities to forge checks and escape from U.S. custody. And after serving five years in prison, he became an American security consultant. Despite its various forms, identity theft has one common goal, that is, to escape from punishments, especially on the Internet. And thieves can make use of it because people decide whether they trust you based on your identity.

Like in real life, identity is not a trivial issue on the Internet. With the boom of internet technologies and all sorts of Internet-based companies, identity is not just a way to differentiate people any more. It affects to what extent people trust each other. One thing people like about the Internet is that it gives them freedom. Freedom of speech, freedom of choosing what they want, freedom of meeting unknown people are just a few examples. Another freedom is that you can be someone else. Currently on the Internet, you can create any username and hide behind it. This name you created is a pseudonym. Unlike the ID in real life, pseudonym is easy to create. In a lot of online games, people often create more than one account and use these accounts to help each other and they or at least one of them can level up faster. And so the Internet is like real life but anyone you meet can be another person tomorrow. Imagine your family members getting changed every morning. This is how the Internet works and also part of the reason why it appeals to us.

However, this freedom is not only given to good people. Thieves can also make use of this freedom. Sometimes, to trap potential victims, the thieves don't even need to get a valid identity. It can be made up on the spot. Therefore, thieves can renew their online identities many times in a short time. The result is a new type of Internet scam: whitewashing (see Section 2.2.4). It's most seen on online stores and e-commerce websites. The thief gets an account using a pseudonym on an online store and tricks a victim into his trap, then he can delete his account in this website and create another account. The victim can report to the online store but they won't know about

the thief's real identity.

It doesn't matter whether you're a seller or a buyer; there are many ways for the thieves to trap you. For instance, if you were a seller, they can go to your page and pretend that they want to buy something from your store. After asking about your PayPal[1] account number, they soon show you a screenshot that they've successfully transferred the money to your PayPal account. But when you check on PayPal, it tells you that you've entered the wrong password too many times today and you can only access your PayPal account after three hours or even longer. The alleged buyer will then urge you to send the product they want. You, the seller, in order to get a better customer evaluation afterwards, won't delay for any minute sending them the product. Later when you can check your PayPal account, there's no record of them paying the money. But you've already shipped your product. Of course you can report this buyer as a malicious user to the website and put the lowest rating on him. However, he would never use this account and can simply get another account. Another example targets the buyer. On an online auction site, you're bidding for a cheap sale but someone else wins the auction. Later the seller contact you saying the winner of this auction gave up and since you're among buyers offering second best price, he's offering you the product. If you agree to take the offer, he would propose a money transfer outside the online auction site. After you transferred the money, he would just disappear and not send you the product and you have no way of getting back your money since you've transferred the money without the supervision of the auction site. And same as in the previous example, you can hardly impose any punishment on him because he can easily get another account on this auction website and start all over again.

According to Javelin Strategy and Research and cited by the U.S. government[2], in 2010, 8.1 million U.S. adults were the victims of identity theft or fraud, with total costs of $37 billion. In 2008, the average out-of-pocket loss due to identity theft was $631 per incident. The number of phishing attacks which attempt to get user credentials on the Internet has greatly increased. And with more and more business, social, government services going online, the frequency of identity theft and fraud are going up.

1.2 Protection against identity threats

One might ask what if those online shopping websites have a strict authentication process so that a person has only one account. While it might be harder for the attackers to launch attacks with fixed identities on the Internet, there are other reasons for not using it. Let alone the freedom of speech, if everyone are exposed to the public, it will be too easy to track one's activities. Like the

[1] http://www.paypal.com
[2] http://www.nist.gov/nstic/NSTIC-Why-We-Need-It.pdf

argument about location based services, no one would feel safe getting services from the Internet if their location information is collected often enough for someone to figure out their path and where they are going. Imposing single ID for everyone helps catching scams but users are in danger of information leaks. And once a thief gets the identity of the victim's, the victim probably cannot be registered anywhere else. The thief can do much more with the identity and probably benefit more than before. Moreover, managing fixed ID is not a trivial issue itself. One of the advantages of pseudonym is that it's easy to create. Fixed identity is usually based on real life credentials and users might not want to give that information just to buy a pair of shoes online. Also if users are releasing these credentials online, with phishing attacks becoming more sophisticated, there's an even higher risk of identity theft and fraud. In 2004 the U.S. government invented the Federal Employee ID Card. This identification card is a smart card and contains information like the employee's name, agency, finger prints, identification numbers and certificate to access different systems. Later on, the REAL ID Act[1] of 2005 proposed national ID cards for the public while privacy advocates have serious reservation regarding Real ID such as the database storing personal information of the whole country getting stolen. Another concern is that it'll be too easy for the government to track its citizens. These concerns stopped the creation of the US national ID. The downside of any form of fixed identity management is its centralization. And probably the best way to protect private information is to not give it to anyone.

And here comes the question: Is there a better way to prevent scams than strict authentication? Identity in itself is not only a set of information; it gets us trust from people and we decide whether to trust another by first knowing his identity. Since identity is the protocol between people, can we make use of it to prevent cheating and thefts?

1.3 Thesis contribution and organization

This thesis begins by introducing online identity and its differences with real life identity. We then present different types of online identities and their benefits and disadvantages. In the next section, background knowledge of Game theory that we used in our analysis will be provided. In the following section, we give results from our Matlab simulation. Finally, we discuss how online identity can be utilized inspired by our results.

[1] http://en.wikipedia.org/wiki/REAL_ID_Act

Chapter 2

Online Identity

This chapter delves into online identities in more details. Online identities can be categorized into three types: fixed identity, pseudonym and social identity.

2.1 Fixed identity

Fixed identity is what we have in real life. Everyone has one and only one identity and no two persons have the same identity. And among all three identity types, fixed identity is the most distinguishing one, mostly because its original purpose was just to distinguish people. As described in the first chapter, identity was invented as fixed identity. It's rarely seen on the Internet but we want to include it in this thesis for comparison purpose. Social security number, driver's license and passport (within a country) can all be considered as fixed identities as well as url addresses which can be seen as identities for resources.

2.1.1 Online fixed identity

It's not easy to find an example of a website using fixed identity. However, there are a few websites that limit the number of accounts a user can register. One of the terms in PayPal's user agreement is no multiple accounts[1]. PayPal has the right to terminate multiple accounts that belong to the same user. Another website, NeoBux[2] which pays you to view the advertisements they show, also has a policy that only one user per IP is allowed in 24 hours and only one account per computer[3]. A user can still register for multiple accounts but he or she will need to have same the number of computers or else it doesn't make any difference having one account or many accounts. This is more like a fixed identity of IP addresses. And the time limit of 24 hours makes it more reasonable

[1] https://www.paypal.com/cgi-bin/webscr?cmd=p/gen/terms-outside
[2] http://www.neobux.com/
[3] http://www.neobux.com/m/a/

that another user of the same computer can still log on using the computer; just have to wait for some time. Also, when Facebook[4] first started, its accounts are based on credentials (emails) issued by educational institutions. Each user had to be enrolled in a school and can be verified with a school email address. And since each student have only one school account, each user could have only one account on Facebook.

2.1.2 Advantages and disadvantages of fixed identity

The advantage of fixed identity is its security, as long as the authentication process can guarantee that a user is really who he claims to be. The disadvantage is its centralization and possible personal information leak. When using fixed identity, a lot of effort is put into building a lengthy authentication process. And usually if a user wants to register for an account, it would take a long time. Authenticating a user also needs many other credentials that are fixed identities. Therefore, an online fixed identity depends on other real life fixed identities. But as more and more technologies evolve to help hackers stealing user information, authentication will become more and more complex for fixed identity. Chaum argues that fixed identity provides one-sided security, protecting service provider from individual users while users' information are left in danger [3]. Their approach to this problem is that individual user creates different digital pseudonyms with different service providers. Transactions are done through a card computer as an intermediate between users and service providers. Thus, only users know about user information, service providers authenticate the card computer with the pseudonym. In this way, the intermediate card computer is like a pseudonym for the user, service providers cannot link pseudonyms.

2.2 Pseudonym

2.2.1 Creation of an online pseudonym

As the most pervasive identity used on the Internet, it's quite easy to create a pseudonym. But we can categorize its creation into two levels based on how much information you have to provide. The easiest ones are those that you can get by giving your name (not necessarily your real name), a name for your account that is visible to other people on the site, a password and maybe a security question. You can get a Gmail account with these information plus a recovery email and your location. Another type is those that you can get with your name, your email address, a password and maybe birth date and gender. Facebook lets you have an account with these information provided but to complete your profile, more information is needed. And it will access the contact list of your email to connect you with your friends on Facebook. It's true that email address is

[4] http://www.facebook.com/

just another basic information but we put this type of pseudonym creation in a different category because the reason that it asks one's email address is to reach to more information: one's contact list, another reason is to ensure that the user is the owner of this email. Adding a mailing address you can have an eBay[5] account.

2.2.2 Organizing online pseudonyms

Most of today's pseudonyms are based on email addresses. Some are based on other pseudonyms. Sites like Mashable[6] and Zoho[7] can be logged in with a Google account. Flickr[8] can let you login with a Facebook account. These are part of the OpenID project[1] which allows people sign into other websites with an established account of OpenID issuing sites [4]. It saves the user a lot of time registering and managing account names and passwords. And for OpenID relying websites, they do not need to put too much effort managing users' identities. Basically, OpenID connects pseudonyms or we can say it reduces pseudonyms. The disadvantage is that if a hacker gets your OpenID he can use your identity on other OpenID relying website. Also the OpenID can be "phished" in more websites. Oh and Jin discuss the limitations of OpenID and show that hijacking attack during the authenticating session is possible [5]. Therefore banking and online trading websites should not jump into OpenID without a second thought.

2.2.3 Advantages and disadvantages of pseudonym

The benefit of using pseudonym is its simplicity. Compared to fixed identity and social identity, pseudonym is the "cheapest". Thus, pseudonym makes it easier for users accessing websites and their services. This also helps the thriving of internet-based companies. Another advantage is that pseudonym protects the real identity of its owner. It's like wearing a mask on the Internet. And many people can enjoy impersonating with multiple pseudonyms. For instance, a user can create one account for his business life and another for social life. And pseudonym can also help in our everyday life. Neubauer and Riedl propose an improved electronic health care architecture which makes use of pseudonym and better protects patients privacy [6]. A good summary of using pseudonym might be the punch line in one of the cartoons published on The New Yorker by Peter Steiner, which showed two dogs sitting in front of a computer and captioned "On the Internet, nobody knows you're a dog" (Figure 2.1).

However, is it really that easy to hide yourself on the Internet? There have been several incidents

[5] http://www.ebay.com/
[6] http://mashable.com/
[7] http://www.zoho.com/
[8] http://www.flickr.com/
[1] http://openid.net/

Fig. 2.1 A dog with a pseudonym

of online pseudonyms being tracked down to the real person. This phenomenon is called "human flesh search engine" during which a big number of netizen searching the web for any information about one specific person who might have done something outrageous. Even if the target is using a pseudonym, he or she will be hunted until the true identity is found. The beginning of this "human flesh search engine" was in 2006 when netizens on an online forum, Tianya[9], were searching for the real identity behind a pseudonym, Poison, who posted a lot of pictures showing off luxurious life style. So maybe the caption in that cartoon mentioned above should be "On the Internet, everyone knows you're a dog eventually". And this "human flesh search" is a way people use the Internet rather than a technical problem. Therefore, pseudonym is not a perfect mask and if a lot of people want to find out about you they can tear down the mask. In [7], Stutzman states that identity information released in a social network community, e.g. photo, politic views, course schedule, could also be used to identify people and should be given more attention. Another way to figure out the real identity behind a pseudonym is to approach the target as someone who can offer him what he wants and just ask about his information. In 2005, a college student named "paula" on the Internet was trying to find someone to write her school paper. A comedy writer responded to her post and offered to write her paper. After talking to "paula", the comedy writer released the real identity of "paula", which brought huge dispute to "paula" and also whether what the comedy writer did was appropriate. Therefore, using pseudonym only gives you limited privacy. Its alleged privacy is achieved by hiding the real name while releasing associated history actions and other

[9] http://www.tianya.cn/

information. One's purchasing history, browsing history and many other history information are shown to other users or even non-users. But if we hide our trace on the Internet and releasing our names it is not such a risky action. Sweeney proposes a k-anonymity model that mixes k users together when a user is submitting his information, so that no one can tell which user in this k users group released this information [8].

Yet there's another privacy issue concerning anonymous systems. In [9], Pashalidis and Meyer present that it is feasible for cooperating organizations to link the transactions of users in an anonymous credential (pseudonym) system. Same type of concern is also mentioned in [10]. Chen and Rahman survey mobile SNAs (social networking applications) privacy designs and conclude that in most of them there is little information provided to the user about what will happen to their personal information in the system once the user submits them. Pashalidis and Mitchell identify limits to pseudonym unlinkability in [11]. In their work, they present how the timing attack, launched by identity issuing and verifying organizations together to figure out which pseudonym belongs to which user, can break pseudonym unlinkability. *Clearly, what pseudonym provides is security against other users but not against the systems that are involved in issuing and verifying the pseudonym.*

The disadvantage of using pseudonym is that it can be easily abused. It provides malicious users with the same amount of safety as the legitimate users. Unlike legitimate users, malicious users are always changing their accounts, making it almost impossible for the website or other users to track their activities. But legitimate users who stay with one or just a few accounts can be traced because as those accounts keep being active within an online community, its reputation will grow and other users will be familiar with them, thus, only legitimate users are exposed to others. This inequity is intrinsic to pseudonyms.

And now more and more people are aware of releasing personal information on the Internet and are more reluctant to interact with strangers. Krasnova et al. define two types of threats that current OSN users are facing, organizational threats and social threats [12]. As a response to organizational threats, users tend to disclose less information about themselves. Regarding to social threats, users tend to consciously control their released information.

Identity theft mentioned in Section 1.1 also happens with pseudonym. Marshall et al. study online identity thefts [13]. In their work, they point out methods of online identity theft: protocol weakness, naive users, malicious software, data acquisition and network impersonation. And pseudonym makes it just easier to steal someone's identity on the Internet. Bonneau et al. show that with just limited information revealed about each user, many properties of the whole social graph can be even reverse engineered [14]. He et al. show in their work that personal attributes can be estimated especially for people who have strong ties (closely related) with other people [15]. Even if people choose not to disclose their private information, it can still be inferred. Furthermore,

these identity theft attacks can be automated. Bilge et al. present how easy it would be to launch automated crawling and identity theft on popular social networks [16].

On the other hand, innocent newcomers are not trusted any more, which is described in Friedman and Resnick's work as the social cost of using cheap pseudonyms [17]. Same problem exists on P2P systems. Adar and Huberman present free riding problem on Gnutella system [18]. P2P systems depend on users sharing their files with others while getting files that other users uploaded. But according to Adar and Huberman, as the community grows and a large number of files on the system are available to everyone, users tend to 'consume' rather than 'produce'. Their experiment show that more than 70% users do not upload files and are taking advantages of other users who are actually contributing to the file sharing system. And they find that free riding is the norm in a large anonymous community like Gnutella. Incentive mechanisms exist for preventing free riding. One is to make use of the downloaded file on user's computer as a replicate for other users to download from. Another one is to make it a rule that one cannot download before he has uploaded enough files. These solutions might reduce free riding but also create their own problems. More importantly, they are vulnerable to whitewashing as pointed out by Feldman et al. [19].

Another type of attack due to pseudonym is Sybil attack where a malicious user creates a number of accounts on a website and use them to outnumber legitimate users. Levine et al. survey solutions to Sybil attack [20]. In their work, they mention several cases where Sybil attack happens. For instance, companies can use Sybil attacks to increase Google PageRank rating in order to boost up their rank in search results. Douceur studies Sybil attacks in large-scale P2P systems and points out that most solutions rely on certain assumptions that are either unjustifiable or unrealizable [21]. Yu et al. leverage social networks to defend against Sybil attacks [22]. Their basic idea is that when there are a lot of Sybil nodes on the social graphs without Sybil nodes, cutting a small number of edges will isolate a large number of Sybil nodes whereas this would not happen on normal social graph. We can see that social network has come to the rescue of pseudonym. Other than Sybil attack, people can maliciously share accounts, for example, friends can share iTunes account to avoid paying for music. And there are even websites that help people to do that. Fortunately, attentions have been paid to this problem. Lysyanskaya has given theoretical constructions for a pseudonym system that discourages people from sharing accounts [23].

Some researchers argue that it is because of pseudonym people think it is fine to be dishonest on the Internet. If this is true, we have a cyclic situation where pseudonym gives us the protection and we, in return, like it more than fixed identity because we can be irresponsible of our words and activities online. To this end, the social networking giant Facebook has a policy that urges users to provide truthful information when registering[10]. "Truthful information" includes no fake name, accurate contact information and no false personal information. And under its Registration and

[10] http://www.facebook.com/terms.php?ref=pf

Account Security section, there are rules that users will not create account for other people, will not create more than one account and will not transfer their accounts. Moreover, Facebook reserves the right to shut down an account if any of the rules were breached. And it did cancel user accounts for above reasons. One of them was the account of "John Swift". John Swift was originally the name of a famous philosopher and writer. The name was used by a blogger on Facebook. Replies from Facebook when "John Swift" questioned about the deletion of his account was that fake accounts were a violation of their Terms of Use. Another account registered under a pseudonym that got taken down was Michael Anti's. A Chinese journalist used the name on published articles, blogs and even Harvard fellowship documents. But it is not the real name of the user. With its real name policy, Facebook is trying to maintain a more accountable Internet environment. And other social networking websites are working toward the same direction. PatientsLikeMe[11], for instance, collects users' information and sell it to medical and drug companies, helping them develop new products. Patients can share his or her medical experience and compare it with other patients suffering the same disease. Obviously, users of PatientsLikeMe wouldn't want fake account. So it also has a rule that users must provide truthful and accurate information. eBay users are also required to not to misrepresent others. Unlike Facebook which is a social network, eBay actually explained with an example why it is imposing this rule. The same example can happen on Facebook even though Facebook is not a trading website. A stranger can introduce himself as one of your close friend or even your family member and ask for money. But ordinary users are often not willing to register with their real name. Regarding to the deletion of "John Swift" Facebook account being deleted, Facebook eventually reactivated his account. No matter how Internet companies are trying to make the dream of a lie-free Internet come true, they cannot convert a single user. Pseudonym is now standing in this awkward situation where some people want to replace it with fixed identity others prefer pseudonym over fixed identity.

Effort has been made to address pseudonym's low accountability. In [24], Penna et al. show that compared to incentive schemes, having reputation associated with pseudonyms has proved to be better for P2P systems. Ford and Strauss propose pseudonym parties that ensure everyone has only one pseudonym while maintaining user's anonymity [25]. Users need to go to "pseudonym parties" on "Pseudonym Day" and get a certificate in person together with a hand stamp to prevent the user form getting a second certificate. The user can then register with pseudonym servers provided by pseudonym server providers. Whenever the user needs to provide his identity to a relying website, the site would redirect the user to the pseudonym server which would authenticate the user. After the user is verified by the pseudonym server, he will return to the relying website and get an account on the website. The user does not have to use the same pseudonym in every relying website. Only the pseudonym server knows about each of pseudonyms a user has on each relying website. In a

[11] http://www.patientslikeme.com/

sense, it's similar to OpenID project in that it links different pseudonyms of one user's and protect these connections. The "pseudonym parties" grant users with one and only one certificate (for at least one year) but it contradicts the convenience of pseudonyms.

But more problems keep emerging because of pseudonyms. For instance, Yokoo et al. study the false-name bids on auction websites [26]. False-name bids are bids that are submitted by a single user but through different accounts. In their study, they show that there is no false-name-proof auction protocol that is Pareto efficient, where everyone gets their optimal benefit. Another example is that on websites like Twitter[12], Google buzz[13] and weibo[14] users can "follow" their friends, celebrities or just strangers. These websites let you post short messages, videos or pictures to those users who are "following" you. Compared to social networks, these websites focus more on sharing contents. This fast sharing trend on website like these has caught the eye of some business people. They register a few account and post useful tips to other users, including news, photos, anything that can get them more "followers". When they've aggregated enough "followers", they can sell these accounts to companies who want to promote themselves. Imagine what if all the "followers" are faked dummy accounts. Since pseudonyms are so easy to get, anyone can create a lot of accounts in a short time. The website itself can actually produce a great number of dummy accounts and use them to boost the website's popularity, thus, it can attract more companies to do promotions on the website. As far as our study concerns, nothing has been done to address this problem from a technical angle.

Pseudonym was first applied on the Internet due to its ease to use and to build. But now it has been creating more and more problems and maybe it's the time to consider reinventing our identity on the the Internet.

2.2.4 Whitewashing

Since pseudonym makes whitewashing possible, we want to introduce the effect of whitewashing in this section. Whitewashing refers to malicious users getting new identities after cheating on others and rejoining the online community to start their attacks all over again. It happens in several different types of websites.

In P2P systems, peers are supposed to upload media objects for other users while getting media they need from other users. Since media sharing is voluntary, users can decide not to upload any file and just get files from other users. These users are recognized by the system as free-riders and penalties are imposed on free-riding. However, the system can only catch free-riding once enough files are downloaded. And because P2P systems are using pseudonyms, free-riders can get away

[12] http://www.twitter.com
[13] http://www.google.com/buzz
[14] http://weibo.com

with a new identity and come back afterwards. This way, a user becomes a whitewasher and can keep free-riding.

On online shopping websites, whitewashers first cheat on some victims and get new identities to rejoin the website and cheat on others. Of course shopping websites can implement strict authentication process and monitor banking information or credit card number, if the whitewasher uses the same bank account or credit card in his "second life", the whitewasher can be soon detected. But some tricks whitewashers have used on other users do not involve any transaction that's under the shopping website's supervision.

Whitewashing in recommendation systems is not rare. It can be used to boost sales of a company's products or to diminish profits of competitors'. There is a business behind this type of whitewashing. Because people would often go to sites like Netflix[15] or IMDB[16] to see comments about the movie they want to watch or check reviews on goodreads[17] or aNobii[18] before buying a book, publishing companies may hire people for flattering comments about their products. Or someone can just do that for fun since pseudonyms on rating sites are free anyway. Obviously, both false positive and false negative comments would harm the credibility of the rating website.

In online games, whitewashing allows malicious users creating multiple accounts and manipulating those accounts to do unreasonable cooperation in order to get to higher levels faster, which is not fair to legitimate users. And more than often, online games only grant points when one beats another. So it will be harder for other users to win in a round of game when more than one accounts belonging to the same user participate in the same round.

2.3 Social identity

As the concept of social network has become popular, social identity was invented. In psychology, social identity refers to an individual's self-perception due to being a member of particular social groups[19]. Websites like Facebook has taken this definition into computer networks and provides a platform that connects people just like how we get to know each other in real life. More interestingly, the Facebook connections a user has do exist in real life, that is, a connection on the social network often indicates that the two users know each other in real life. A lot of applications have been developed to leverage these connections created on social networks. They can also be utilized for identifying people. And the social identity we are talking about is not just a Facebook profile but a digital identifier generated based on the social graphs that SNSs like Facebook have collected.

[15] http://www.netflix.com
[16] http://www.imdb.com/
[17] http://www.goodreads.com/
[18] http://www.anobii.com/
[19] http://en.wikipedia.org/wiki/Social_identity

2.3.1 Making use of connections

Facebook connections are not the only virtual connections. Email contact book, MSN contacts, RSS subscription and even bookmarks on a user's browser can be translated as connections. They all stand for either one-way or two-way relations. An early but popular example is when Google first launched Gmail[20], a user can only get an account if someone invited him. But if someone does not have an email in the first place, he cannot receive an invitation. And if a non-famous company starts a service based on invitations, it might discourage users from using this service. Another example is importing the user's friend list from MSN and email account and directly connecting the user's friends on the new website. Thus, the user would visit the website more often because his friends are there too. For online trading websites, they can investigate into users trading history and rate each user's credibility to make their sites safer and more accountable. Brainard et al. propose using friend information to authenticate a user [27]. When the primary authentication, a password or a token, is not available to the user, an emergency authentication can be used. By emergency authentication, a "helper" can vouch for another user, the "asker", with his primary authentication. Soleymani and Maheswaran has proposed a scheme that puts Brainard's fourth factor in use when authenticating users in the mobile social context [28]. Essentially, a mobile phone is something that's carried by the user himself and thus can be used to verify that this user is a friend with a claimed friend. For instance, friends will have phone calls or are able to connect through Bluetooth when they meet. Therefore, these phone calls and Bluetooth connections can be used to validate the relationship between the "helper" and the "asker". Other than individual connections, the topology of a social graph is also used to defend against Sybil attacks. In [29], Viswanath et al. investigate into existing social network-based Sybil defenses. Despite considerable differences, all existing Sybil defense schemes depend on community detection and are based on assumptions that Sybils can only form a few connections to non-Sybils and that the presence of Sybils lead to misbehavior. Viswanath et al. discover that the reliance on community detection makes existing Sybil defenses vulnerable. And sometimes, legitimate users do create multiple accounts. Our approach, on the other hand, does use social graph to generate social identity but we are not detecting "whitewashers". We focus on the history of a user instead of trying to mining everything out from the structure of the social graph. Regarding identity theft we mention at the beginning, Jin et al. have characterized the behaviors of identity clone attacks and devised ways to detect suspicious identities in [30]. Their approaches are more active than existing third-party application on OSN that help verifying a user by taking into account the friend lists and attributes of a profile when calculating profile similarity. However, in case of "whitewashing", where users are not necessarily building profiles and adding friends, detecting "whitewashers" might not be

[20] http://www.gmail.com

effective. Because authentic users may have poorly maintained profiles which makes it harder to distinguish between legitimate users and "whitewashers" just from the attributes and friends list. But once a profile is detect as being a faked one on an OSN, any account related to this profile on other networks can be labeled as "whitewashers" and this would prevent "whitewashing". Another evidence of social network's accountability is provided in [31]. Mislove et al. show how to infer attributes of other users given only the information of some users on the OSN.

Other than its accountability, social network has a nice property that people can find each other faster. In [32], Watts et al. offer a explanation for social network's searchability. Mislove et al. give the structural measurements of online social networks [33]. They find that online social network has special advantages in disseminating information and inferring trust among users. Subramani and Rajagopalan design a framework for viral-marketing, where interested consumers can tell each other about the product, based on online social network [34]. Leskovec et al. even use online social network to predict people's attitudes towards other based their connections with their friends [35]. Kimball and Rheingold present how online social network can help big organizations to get organized and communicate efficiently and hence reduce their internal friction [36]. Mankoff et al. try to encourage users to reduce their ecological footprints and it is expected to have a large impact on users [37]. Cachia et al. claim that online social network can foster creativity, reflect changes in social behavior and accumulate collaborative intelligence [38].

2.3.2 Generating social identity

Using connections to generate identity for everyone is different from above applications in that identifying someone is not a local issue but a global one. Maheswaran et al propose a framework that generates social identities for users and relying parties can use social identities for authenticating users [39]. Each user in social network can be considered as a node and their connections are links between the nodes. Thus social network can be viewed as a connected graph. Social identities are generated from the view of a landmark server, which is connected to strategically chosen nodes in a social graph. From the perspective of the landmark server, we can divide nodes into sets based on their levels by running BFS on the social graph. A user's social identity consists of a set of edge-disjoint paths from the landmark to the user. Only the user can access his social identity. We can bind user's social identity with the user using a user-centric identity selection framework as shown in Figure 2.2 [39]. The user can register with a social identity provider using a public-private key pair. Every time the user needs to be authenticated, the identity selection framework will ask for the minimum requirement and it can then interact with the social identity provider and request for a social identity. The social identity provider then start its verifying process to make sure that the user's social identity is still unchanged or calculating the latest social identity due to newly created or deleted connections. Heartbeat messages are sent regularly throughout the social graph

to detect new or lost connections. The user now has his latest social identity. If he needs to prove his identity to any party, he can give this social identity to them and they can verify this social identity with the social identity provider. In Figure 2.2, users establish their social identity through step 1, 2 and 3. Relying parties can verify the social identities submitted by users through step 4 and 5.

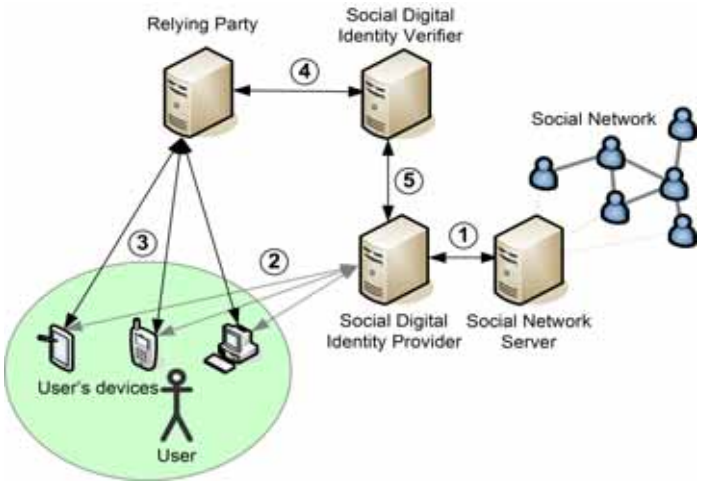


Fig. 2.2 An example deployment scenario for social digital identities

2.3.3 Properties of social identity

As this thesis is not dedicated to generating social identity, we describe properties of social identity that make it different from fixed identity and pseudonym. Social identity is based on connections one has with other people not credentials or tokens provided by the user. Users do not need to provide their critical real life credentials like social security number, driver license or household registration. You don't even need to register your real name. Social graphs are provided by social networks. Only the social network that's providing the social graph knows about the topology changes on the graph. It'll be harder for malicious users to forge identities. Regarding to user privacy, when the user provides a social identity to a third party, the third party cannot crack the information out of the social identity, it can only send the social identity to a social identity provider and get answers about its validity. And so, social identity protects user's real identity better than fixed identity. Since only registered users can issue social identity for themselves and users register with the social network by public-private key pair, one cannot ask the social identity provider to issue an identity for other users. A user can have multiple accounts, for instance, one for work, another for personal relations. But to be authenticated by other parties, different social identities will be generated for each account and verified. The generated social identity can be

used anywhere when the user needs to provide identifications, so it's a portable identity, unlike the case of pseudonym where each user probably has multiple different identities all over the Internet. On the other hand, websites do not have to build their own identity management system. Social identity provider and identity selection frameworks will do the job for them. Thus, the Internet will become a virtual community. Then users can decide whether to trust someone just like in real life in contrast to pseudonym with which new comers are likely not to be trusted.

The most essential part about social identity is the interactions between social networks, social identity provider and identity selection frameworks. Social networks provide the data that is needed to generate a social identity for a user. There can be multiple social networks feeding data to a social identity provider and they need to guarantee that the social graph is secured and will not be released to any third party, even the government if not necessary. Also, in any case, user himself is the one who asks for the issuing of his own social identity and decides to whom to release his social identity. No one should be given social identity of others'. Social identity is user-centric. It puts user in charge of his credentials and personal information. Other people can only verify a social identity through identity selection framework. Social identity provider only accepts messages from identity selection frameworks. This is to prevent hackers from making up fake messages and get the social identities of others.

2.3.4 Application of social identity

As pointed out in [40], the business world has adopted the concept of social networking and there is a growing need for a unified digital identity resource. Social identity is not merely a name but reputation of the owner of the identity. It hides the real identity of the owner's from organizations and provides organizations with a convenient way to authenticate the owner. Therefore, social identity protects both parties. Because social identity is based on people's relations, we can design better trust models if we categorize people's relations into different trust levels. In this way, we can always improve the online community. Jennings and Finkelstein propose a trust model that can help in constructing community-aware identity management systems [41]. Agarwal et al. present the long tail distribution of social networks: most people have few contacts and few people have a lot contacts. They show how social graph can be used to help find the "familiar strangers", people who don't know each other but exhibit similarities [42]. And if the virtual community is more accountable, it can enlarge the extent to which we can rely on the Internet. For example, we can do groceries online, first send the shopping list to the store online and once the store verified that you're a valid resident in its service area through social identity service provider, it can then prepare your grocery and you can pick it up as you finished a whole day of work. Or you want to know about your neighbors better but you don't have the time to social with them, you can interact with them online and organize events together. A lot of public services can be redesigned

to our needs. When the city is planning to build another metro line, it can launch a survey online and people living around candidate metro lines can express their opinions. It's not that we cannot do this using fixed identity or pseudonym but that social identity does offer more trustworthy information without compromising any party's privacy and security. The Internet and real world are still separated now but social identity can bring them closer.

2.4 Summary

Identity is never a trivial issue. It decides how we interact with each other and how much we can trust other people online. With an ideal identity system, the user can choose to reveal or hide his information and to whom to give his identity as well as to what extent he wants to publish his own identity. Identity should not be a burden but an access to all sorts of services. It can enhance security of organizations and companies without compromising users' privacy.

Chapter 3

Game Theory

In this chapter, we introduce evolutionary game theory and the concept of evolutionary stable strategy. Later we present the four strategies we use in our study.

3.1 An introduction to game theory

A game is composed of players, strategies and payoffs. Game theory studies how players following different strategies (or same strategies) interact with each other and what the consequences are for each of them. The process of making a decision and the following interaction between the players is the game. The interest of game theory mainly focuses on equilibria under different scenarios. Below we introduce the basics of game theory with regard to its application in various fields.

3.1.1 Components of game theory

Players are the entities that behave under a defined strategy or a plan. A player is not necessarily a person. Organizations, companies, computers or cell-phones can all be considered as players. Players are supposed to make decisions about their "moves" against others. They can either cooperate with or defect on other players in a game. But their purpose is to maximize their benefit from the game or minimize their loss. It's possible that a player may not aim at maximizing his profit but simply follow a rule he believes as the best choice. However, those who make optimal decisions, which we call rational players, are likely to stay in a game and others who make suboptimal decisions might be eliminated eventually.

Strategies are rules describing what a player would do in all possible situations. It can be a simple yes or no choice or a matter of how much to bid in an auction. There are two principles regarding how players design their strategies. One is that players have well-defined preferences over the possible outcome of their decisions or "moves". Another is that players can take into account

other players' decisions. A strategy is not just an action; it dictates what actions a player will take in response to other players' actions.

Payoff (or utility) is the net benefit a player gets after a round of game. It can be positive, negative or zero. Payoff can be concrete objects, such as, number of products or money as well as abstract objects, like trust and leadership. But in order to study player's behaviors and profit in a game, payoff has to be quantified. Payoff of a player takes into account his benefit and loss or cost of his decisions in a game.

Equilibrium is a set of strategies for all players and everyone would stick to his current strategy. Game theory has defined several different equilibriums based on what the outcome they achieve for players. The most famous one is Nash equilibrium which yields optimal payoff for every player given the strategies of other players. It is a state where no player in a game would gain a better payoff by changing his strategy while others stay unchanged. Other equilibriums include correlated equilibrium where no player would deviate from a recommended strategy from a third party, without knowing what the third party recommends to the other player and symmetric equilibrium where everyone employs the same strategy. Various other equilibriums also exist. Despite its various forms, equilibrium does not happen because players are rational, but in an equilibrium players act rationally. Classical game theory mainly focuses on what the equilibria are there in a game, how to select a better equilibrium from all equilibria and how to get to a desired equilibrium, and more practically, how to make decisions.

3.1.2 An example of game

Let's have a look at how to play a game in game theory. Taking the game of Hawk and Dove for an example, two drivers are diving towards each other on the same lane. They can either continue driving forward or stop and turn back. If they both choose to keep driving, a car crash will happen and they will both be injured. If one of them decides to turn back, he cannot get to his destination and the other driver gets to drive safely towards his destination. If they both stop, since there's no one coming towards them, they end up staying there. We can think of the distance between the drivers and their destinations as payoffs. Thus, we can describe the game using the following matrix.

Driver2 / Driver1	Keep driving	Turn back
Keep driving	Crash, Crash	Going through, Going back
Turn back	Going back, Going through	Stop, Stop

Table 3.1 Driver's Game

The payoffs for the two drivers are: (Crash, Crash), (Going through, Going back), (Going back,

Going through) and (Stop, Stop) are the payoffs for the two drivers. The first entry is the payoff for driver1 on the leftmost column, and the second one is the payoff for driver2 at the topmost row. Further, we can quantify the payoffs and we get the following payoff matrix. The payoff values do not have to be 0, ±1 and 10, it only means that the driver has a preference over the four outcomes: going through than stop, stop than going back and going back than crash.

Driver1 \ Driver2	Keep driving	Turn back
Keep driving	-10, -10	+1, -1
Turn back	-1, +1	0, 0

Table 3.2 Hawk and Dove

A payoff matrix like this is a standard way to describe a game in game theory. Another presentation is the extensive form. Each player has to choose between keep driving or turn back, their strategy is a choice between the two options. Since they want to avoid crashing, turning back is a safer choice than keep driving. A more risky driver may decide to keep driving as long as no car is coming from the other direction. If the driver assumes that the other driver is reasonable enough to turn back before he does, he can just keep driving. Therefore, there are two Nash equilibriums in this game, (+1, -1) and (-1, +1). Recall Nash equilibrium's definition, at (+1, -1), if driver 1 changes his strategy to turning back, he gets 0 payoff which is less than +1 and if driver 2 changes his strategy to keep driving, he gets -10 which is also less than -1. Thus, no one gets more by changing his strategy unilaterally.

3.1.3 Applications of game theory

Game theory originated from analyzing how one can do better in a competition and has now been applied to a wide range of problems. This section tries to categorize applications of game theory.

How to make decisions

This was the initial purpose when game theory was invented. From economic to philosophy, game theory has been used to assist decision makers and scientists analyzing their situations and getting a clearer picture of how people make decisions. For instance, a well-known example is that when companies are setting prices. If companies cooperate with each other, they can set the price high, which is profitable to all of them. If they try to compete with each other or just one of them decides to lower the price, the price will be lower. And eventually, other companies will have to lower their prices. This situation has been studied using the Prisoner's Dilemma. Another example is when a company wants to get into an already well developed market and another company is the dominating company or a monopolist. The entrant company needs to choose between getting into

the market or staying out of it. The dominating company has the choice of accepting the entrant company or competing with it. There are many applications in philosophy too. Bruin presents a survey of different philosophy branches, which game theory has been involved in [43]. On topics like epistemological dependence, liberalism and efficiency, morality and rationality, and many more, game theory is used as a tool for studying human logic.

How to achieve equilibrium

Equilibrium has attracted interests from game theorists since long ago. What is the meaning of equilibrium? Equilibriums are possible agreements that a game would finally reach at. And studying equilibrium may help us design better schemes for policy makers and engineers. For example, in wireless cellular systems a base station is responsible for sending and receiving data to and from mobile phones. But at the edge of a base station's coverage, signal strength may be very low and if we can make mobile phones help each other relaying data, data rate of each mobile phone can be improved. But if mobile devices are selfish, they would not agree to relay data. Engineers have used game theory models to investigate how punishments and awards can be imposed to encourage individual device cooperate with each other. For policy makers, game theory can help decide rules that can turn undesired outcome to a more desirable outcome. Webster presents case studies showing that studying equilibrium can improve common laws to be more efficient [44].

Modeling and prediction

Game theory is a good tool for predicting outcomes of situations where multiple interests are involved. Bruce Bueno de Mesquita, a consultant to the CIA and the Department of Defense and also a politic scientist, actually predicted what Iran would do in ten years using game theory. And his model is 90% accurate while the government experts are wrong. Goodwin has introduced more ways in which game theory is engaged in making predictions [45]. Amsel, Pilpel and Marshal developed a game theory model to better explain Obsessive Compulsive Disorder (OCD) [46]. The advantage of game theory is that it can combine different logics interacting with each other and hence provides a way to mix together cooperating or conflicting parties. Moreover, evolutionary game theory, which we make use of in this thesis, is especially good at modeling on-going strategic interactions between large population of players. The following section gives an introduction to evolutionary game theory.

3.2 Evolutionary game theory

Evolutionary game theory differs from classical game theory in that it focuses on games played by groups of players, not any individual player. It first appeared in [47] where Fisher observed

dependence between individual's expected number of grandchildren and the ratio of males and females in the mammals' population. In 1961, R. C. Lewontin brought evolutionary game theory into evolutionary biology. Later, evolutionary game theory has become popular in the fields of economics, sociology and anthropology. Although its applications span a wide range, the key properties remained that it's a dynamic theory. And, with an emphasis on interactions of groups of players, it reveals more than just equilibria of games.

Before introducing details of evolutionary games, we need to take a look at its assumptions and how it completes game theory in fields that classical game theory could not achieve. As argued in [48], one of the assumptions of classical game theory is that players are rational. But the truth is human are not always perfectly rational. Most people simple learn as they make mistakes. Classical game theory doesn't give a chance for players to learn. That is, behaviors that lead to rational equilibrium may not be rational themselves. Another contradiction is that when a player deviates from his equilibrium strategy, classical game theory tends to believe that he did want to move towards equilibrium but was interfered by some other forces or just a slip of hand. Evolutionary game theory believes that the player is using a non-equilibrium strategy and thus, reasons differently about the game. Rationality is better at playing against rational players. When playing with irrational opponents, it's better if we can change our strategy along the way and avoid being exploited. Evolutionary game theory let players change their strategies by repeating the same game over and over. Below, we give a more detailed introduction of how evolutionary games evolve.

3.2.1 Evolutionary game theory setting

Modeling with evolutionary game theory has two steps: defining initial population composition and updating offspring of each strategy as the game continues.

Population composition

In evolutionary games, it's not arbitrary numbers of players that are playing the game, but a population with defined proportions of players using different strategies. Each individual player behaves just like in classical game theory but has limited effects on how much payoff others are getting at the end. Because the game is played repeatedly, an individual player meets different players in each round, thus, his payoff is continuous. To model this dynamic interaction, we assume the whole population of players is 1. According to how the game is played, we further slice the whole population into proportions of different types based on their strategies or their states. As the evolution goes on, players are changing strategies and the distribution of players using different strategies will also change. Thus, we have a full composition of the population at any time of the

game and this will show us how the population is changing during the game.

Like in classical game theory, we have a payoff scheme for the game. However, we do not focus on individual player's loss or benefit. After each round of the game, we calculate the payoff of each group. The payoff matrix is not different from that in classical game theory.

Replication dynamics

Since evolutionary game theory deals with "ongoing" games, a game is repeatedly played and players can change their strategies [49]. What we are interested in is which strategy will survive, or rather, how a particular group of players win over others throughout an on-going repeated game. And players care more about their long-term benefit rather than payoff of a single round. When a player is making his decision, he will take into consideration how his action in this round will affect other players' future moves and his own payoff in the following rounds. However, the payoff one gets in the future is not as immediate as the payoff one gets now. For instance, during a business negotiation, the market price may fluctuate or there may be other chances that are missed due to the current negotiation. In essence, "time is money" does reflect the truth. And in game theory, the nature, e.g. whether it will rain tomorrow, is sometimes considered as a factor that can have a decisive effect on a player's decision. Therefore, in ongoing games we need a way to evaluate future payoffs. This is referred to as "discounting" in [50].

Players would always like to get his investment back from the game as soon as possible and may see future payoffs less important than current payoffs. To integrate the effect of discounting into evolutionary games, we apply a discount factor when calculating total payoffs of players. Although, it should be noted that "discounting" is not only applied to evolutionary game theory, other branches of game theory where the game has multiple stages also have this concept. Under evolutionary game theory, a player plays a game infinite times, thus, we need a factor to reflect the time difference of each round. On the other hand, it's not sure how long a player will stay in a game, especially in our case of online interactions, users can decide they do not need the service of a website and just never show up on that site. All of these considerations are summarized with the discount factor, which we name as w. Suppose a player will get a payoff of a in each round, his expected total payoff will be calculated as:

$$\text{Expected Total payoff} = a + aw + aw^2 + aw^3 + ...$$

Define,

$$s \equiv 1 + w + w^2 + w^3 + ...$$

then,

$$s = 1 + ws$$
$$s - ws = 1$$
$$s = \frac{1}{1-w}$$

and we have,

$$\text{Expected Total payoff} = \frac{a}{1-w}$$

In evolutionary game theory, expected payoff is referred to as "fitness", which is borrowed from Darwinian fitness.

Another dynamic property of evolutionary game theory is that the population composition changes, which is the second step mentioned at the beginning of this subsection. Evolution in evolutionary game theory goes on just as in biology. It is pointed out in [48] that evolutionary game theory borrowed the evolution from biology. Genes are the players in biology and their hosts' behavior are their strategies. The payoff of a gene is the number of offsprings carrying that gene. However, players in biology cannot change their strategies. But we need some way to allow human players to change their strategies. And this is done through the replicator dynamic, where population of a specific strategy will be updated according to its total payoff after each "generation". Generation is an arbitrary number of rounds of the repeated game. The number of offspring of a strategy is updated as follow:

$$p_i^{t+1} = p_i^t \times \frac{V_i^t}{V^t}$$

We use the presentation of the replication dynamic in [51], where p_i^t is the population of strategy s_i in t^{th} generation, p_i^{t+1} is the population of strategy s_i in $(t+1)^{th}$ generation and V^t is the average payoff of the whole population and V_i^t is the average payoff of the population playing strategy s_i.

Evolutionary stable strategy

Like classical game theory tries to solve and analyze equilibrium, evolutionary game theory studies *evolutionary stable strategy* (ESS). An ESS is equivalent to Nash equilibrium in classical game theory. It states that a strategy is resistant to invasion of mutant strategy. That is to say, when an ESS is established in the population, if a small number of players deviate to another strategy, the mutant strategy will be eliminated. The mutant strategy is eliminated due to "natural selection" — the replication dynamic. And since the evolution is based on fitness of each individual strategy, an ESS needs to get more payoff than other strategies when played together. As in our case, we

want to see how normal users can eliminate whitewashers. That is, staying with one identity can gain the user a higher fitness than being a whitewasher.

3.2.2 An example of an evolutionary game

Here, we would like to use part of [52] as a simple example showing the analysis of evolutionary game theory. In Axelrod and Hamilton's work, they investigate different strategies playing Prisoner's Dilemma in a evolutionary context. The payoff matrix is given below.

Player1 \ Player2	Cooperate	Defect
Cooperate	3,3	0,5
Defect	5,0	1,1

Table 3.3 Prisoner's Dilemma

One of the strategies is Tit-for-Tat: cooperate on the first move and then for each round after the first round, copy the opponent's last move. The other one is AllD: always defect on the opponent. Let w be the discount factor. When Tit-for-Tat is playing against another Tit-for-Tat, it gives the player a payoff of 3 in each round. And the expected payoff is:

$$V(tft|tft) = 3 + 3w + 3w^2 + ... = \frac{3}{1-w}$$

When Tit-for-Tat is playing against AllD, the AllD player gets 5 in the first round and 1 in every round after. Therefore, it's expected payoff is:

$$V(AllD|tft) = 5 + w + w^2 + ... = 5 + \frac{w}{1-w}$$

Hence, in order for Tit-for-Tat to be evolutionary stable against AllD:

$$V(tft|tft) \geq V(AllD|tft)$$

which can be induced to:

$$w \geq \frac{1}{2}$$

However, this analysis does not include a population. In the next chapter we will introduce our models where games are played by a population of players.

3.3 The game and strategies used in our study

In this section, we introduce the game and strategies we used in our model.

3.3.1 The game in our model

We use the donating game in [51]. In each round, users are paired up randomly. In each pair, one player is selected as the donor and the other as recipient. The donor decides whether he wants to cooperate with the recipient and if he does, he has to pay a cost of cooperation and gets a payoff of $-c$. The recipient gets b and $b > c$. If the donor decides not to cooperate, both get 0.

Donor \ Recipient	Acept
Cooperate	-c,b
Defect	0,0

Table 3.4 The Donating Game

This seemingly simple game is similar to online interactions since it usually amounts to a user getting the benefit and the other offering the benefit. Also, no one is always a donor or always a recipient. People tend to benefit from the Internet by interacting with many users. A donor can expect getting back his previous investments by becoming a recipient next time.

The strategy of whitewashers is to never cooperate and change their identities every time after a defection. But legitimate users (which we will refer to as "discriminators") can have different strategies. We analyzed four strategies for discriminators and they are introduced in the following sections.

3.3.2 Image score

Image score is based on the concept of indirect reciprocity which was first designed into a strategy by Nowak and Sigmund in [53]. Each player has an image score to indicate whether he cooperated or defected last time as a donor. Discriminators only cooperate with those who cooperated with others, meaning, whose image score is 1. When paired up with a player who has an image score of 0, the discriminator will defect. The whitewashers will defect on everyone. Therefore, the number of whitewashers with image score of 1 will only decrease, turning into whitewashers with image score of 0.

3.3.3 Pay your dues

In [17], Friedman and Resnick proposed a "pay your dues"(PYD) strategy which rewards cooperating users instead of punishing non-cooperating ones. Under PYD, when a new player is joining in, he has to pay "membership fees" to the community by cooperating with existing players. When a new player is paired up with a veteran player, the veteran player get to defect on the new player. Thus, the veteran will get b as his payoff and the new player pays c as the "due". A newcomer

has to pay in total b amount of dues before he can be considered as a veteran player. We use a flag to indicate whether a player has paid off all his dues. A flag of 1 means the player has paid all his dues and a flag of 0 means the player still needs to pay his dues. A discriminator will pay dues only to veteran players but a whitewasher would want to pay dues to everyone and become a veteran as soon as possible so that he can defect on others.

3.3.4 Pavlov

Pavlov cooperates in the first round and only cooperates with his opponent if his opponent did the same thing with him in the previous round. If a donor who cooperated in the last round meets a recipient who also cooperated in the last round, the donor would cooperate with the recipient. If he meets a recipient who didn't cooperate in the last round, the donor would defect on the recipient. But if the donor himself didn't cooperate in the last round, he would defect on those recipients who cooperated in the last round and cooperate with those recipients who defected in the last round. Therefore, we need an indicator of whether a player cooperated or defected in the last round. An indicator of 1 implies that the player has cooperated in the last round and 0 means the player defected in the last round. Whether a discriminator cooperates with his recipient depends only on whether the indicators of the donor's and the recipient's are the same. But whitewashers never cooperate like with other strategies, no matter his indicator is the same or not with his opponent's.

3.3.5 Tit-for-2-tat

Tit-For-2-Tat (TF2T) is a strategy with which players cooperate at the beginning and defect on those who has defected twice consecutively. This is like the image score strategy but we only toggle the image score of a player's from 1 to 0 when the player has defected two times in a row as a donor. Discriminators are being lenient and whitewashers can get more benefits but discriminators can gain from cooperating with themselves because when a discriminator defects once, others will still cooperate with him. We want to include TF2T in our study because Nowak and Sigmund show that a more forgiving type of image score gains more payoff than the normal image score [54].

Chapter 4

Modeling Different Strategies

As mentioned in 3.3, we have four strategies and for each of them we want to study the performance of fixed identity, pseudonym and social identity. Hence, we have 12 cases in total. Before we set off building models for various cases, we need to get a clear idea of how the three identity types are different from each other. Using fixed identity means a player has one and only one unique identity and this identity can be canceled, but if the same player comes back to the game, he will have to create the identical identity. Throughout the game, a player owns only one identity and no two player has the same identity. With pseudonym, any player can create an identity for himself at any time but one pseudonym is used by only one player. Creation of a new pseudonym costs nothing. Social identity uses the connections a player has with other players to identify a player. To participate in a game, a player needs to connect to at least one existing player on the network. Creating a new social identity is also free.

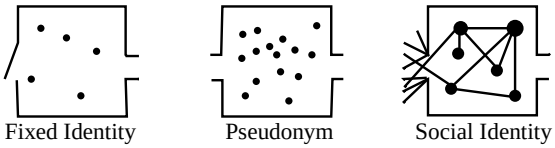


Fig. 4.1 Identity types differences

Figure 4.1 illustrates differences between the three identity types. Every identity type provides us with an arena to interact with other people. A player can only enter the fixed identity's arena once but can leave as he wants. Fixed identity has a one-time entrance and a free exit. Pseudonym's arena does not regulate the come and go of any player so it has a free entrance and a free exit. Social identity connects players while in the other two arenas, players are isolated. A player can

only enter the arena by connecting to someone who is already in the arena.

In section 4.1, we provide the common settings of our modeling of different strategies. Section 4.2, 4.3, 4.4 and 4.5 explain the modeling of the four strategies mentioned in section 3.3.

4.1 Basics of models

4.1.1 Initial population

Like discussed in 3.3, for each strategy we need to label the players with 1 or 0. In image score, the label denotes whether the player has cooperated in the previous round. In PYD, the label stands for whether a player has paid all his dues. In Pavlov, the label again represents the player's last move. For TF2T, the label shows whether a player has defected twice consecutively. Therefore, we have the following population composition. For all four strategies and with all identity types, the whole population is considered as 1, x is the population of discriminators, legitimate players who are not malicious whitewashers, and y is the population of whitewashers. The population of discriminators that have been labeled with 0 is x_0, and x_1 are the discriminators that have been labeled with 1. The populaiton of whitewashers that have been labeled with 0 is y_0, and y_1 are the whitewashers that have been labeled with 1. Thus, we have $1 = x + y = x_0 + x_1 + y_0 + y_1$. And we divide each type of player equally so that half of x_0 will be donors and the other half will be recipients. Same for x_1, y_0 and y_1.

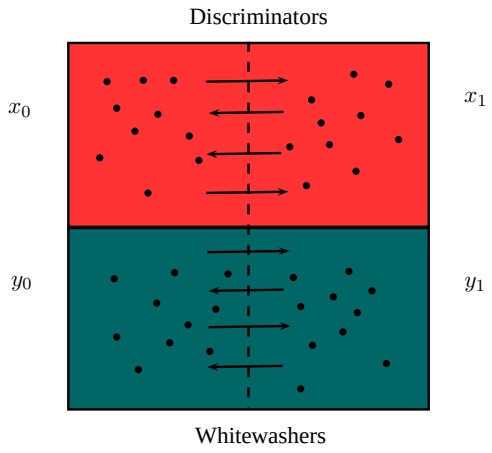


Fig. 4.2 Population setup

Figure 4.2 shows that, at the very beginning of the game, there are equal number of each type of players, and after each round x_0 can become x_1 and x_1 can become x_0, y_1 can become y_0, only in PYD y_0 can become y_1.

4.1.2 Interaction dynamic

For all three identity types under image score, Pavlov and TF2T, y_0 has no way of becoming y_1 but under PYD if a y_0 has paid all his dues, he can become a y_1. For fixed identity, y_1 cannot become y_0 because fixed identity forbids changing identities. For pseudonym, y_1 can always come back to the population as a y_0 and same for social identity. However, in social identity, the label of 1 or 0 is given by others who have interacted with the player before, not the player himself. In pseudonym, this indicator of whether a player cooperated is labeled by himself or the "system" but it will be very easy for a whitewasher to create a dummy user and simply cooperate with himself. In social identity, a player can only interact with players he is in the reach of, for instance, his opponent is within 3 hops with him, or within 3 hops of the players he has interacted with before. Thus, creating a dummy user does not give a whitewasher a lot of potential targets. All of these is reflected as a lower credibility of the labels in pseudonym. That is, a discriminators seeing a label of 1 can choose not to believe it and still defect on his opponent. But with social identity, the player will believe his opponent with a label of 1.

The game continues on generation by generation. Each generation consists of fixed number of rounds of games, depends on the value of w, and x and y stay the same during one generation. Only x_0, x_1, y_0 and y_1 change. At the end of each generation, we calculate the average payoff discriminators and whitewashers get. And then we calculate average total payoff of the whole population during the past generation. Discriminator's and whitewasher's population in the next generation is computed as follow:

$$x' = x \times \frac{V_{disc}}{V}$$
$$y' = 1 - x'$$

where V_{disc} is discriminators' average payoff sum through the last generation and V is the whole population's average payoff sum in the last generation. And x is the fraction of discriminators in the population in last generation and y is that of whitewashers', x' is discriminators in next generation and y' for whitewashers in next generation. At the beginning of next generation, we set x_0 and x_1 back to $\frac{x}{2}$ no matter what the ratio of $\frac{x_0}{x_1}$ was at the end of last generation. Here we can see that if one type of player's average payoff sum of a generation is always higher than other strategies, more and more players will use this strategy, and so it becomes ESS.

4.1.3 Parameters and representation

In our model, we need two probabilities: p and q. Probability p is how often a discriminator cooperate with an unknown recipient whose label is also unknown. Probability q represents how

often a discriminator knows the recipient's label. We need these two probabilities because users on the Internet should have options whether to release their interaction record or not. Revealing a positive score would help build a user's reputation but we should put users themselves in charge of their own information. The representation we use for population changes can be explained as follows:

$$\text{which type of players when selected as} \begin{cases} \text{donor} \\ \text{recipient} \end{cases} \text{ when}$$

$$\text{paired up with} \begin{cases} x_0 \rightarrow \text{have how much possibility} \\ x_1 \rightarrow \text{have how much possibility} \\ y_0 \rightarrow \text{have how much possibility} \\ y_1 \rightarrow \text{have how much possibility} \end{cases}$$

$$\Rightarrow \text{will} \begin{cases} \text{stay as the same type of players} \\ \text{become what type of players} \end{cases} \text{ in the next round}$$

And the representation we use for measuring payoff is explained below:

$$\text{which type of players when selected as} \begin{cases} \text{donor} \\ \text{recipient} \end{cases} \text{ when}$$

$$\text{paired up with} \begin{cases} x_0 \rightarrow \text{have how much possibility} \\ x_1 \rightarrow \text{have how much possibility} \\ y_0 \rightarrow \text{have how much possibility} \\ y_1 \rightarrow \text{have how much possibility} \end{cases} \Rightarrow \text{ will get } \begin{cases} \text{-c} \\ \text{b} \end{cases} \text{ in the next round}$$

4.2 Image score

As described in section 3.3.2, every player has an image score indicating whether he cooperated last time as a donor. Donor decides whether he's cooperating based on the image score of the recipient's. Image score has values of 1 or 0. An image score of 1 means the player cooperated with his recipient last time when he was selected as a donor. And an image score of 0 means he didn't cooperate last time. Donors only cooperate with those recipients who have image score of 1, otherwise donors do not cooperate.

4.2.1 Fixed identity

We borrowed the equation from [52], and setting their $q = 1$ and $p = 0$ gives us the expected payoff of discriminators and defectors in k^{th} round:

$$P_{disc}(k) = bx(\frac{1}{2})^k - b(\frac{1}{2})^k + \frac{1}{2}(b-c)(\frac{1+x}{2})^{k-1} \tag{4.1}$$

$$P_{ww}(k) = \frac{1}{2}bx(\frac{1}{2})^{k-1} \tag{4.2}$$

If we sum up their generation total payoff, we get:

$$P_{disc} = \frac{b(1-(\frac{w}{2})^r)(x-1)}{2-w} + \frac{b-c}{2} \times \frac{1-(\frac{1+x}{2}w)^r}{1-(\frac{1+x}{2}w)} \tag{4.3}$$

$$P_{ww} = \frac{bx(1-(\frac{w}{2})^r)}{2-w} \tag{4.4}$$

and now we can calculate average payoff sum in the whole population:

$$P_T = xP_{disc} + yP_{ww} = x \times \frac{b-c}{2} \times \frac{1-(\frac{1+x}{2}w)^r}{1-\frac{1+x}{2}w} \tag{4.5}$$

4.2.2 Pseudonym

We modified Feldman and Chuang's work from [51]. It should be noted that [51] assumed that y_1's image score is not visible to others but we argue that since they have an image score of 1, they must have not changed identities yet. Their image score should be visible to others by the same probability of q like discriminators. Besides, players can decide not to trust an image score of 1 with pseudonym. And this probability is y – *the population of whitewashers.*

Population changes

We need to first figure out the population frequency changes, that is, what player will become x_0, x_1, y_0, y_1 in the next round. And so we analyze all the possibilities of who's paired up with whom and who's selected as donor and recipient. No matter how image score flips between 0 and 1, discriminators are discriminators and defectors are defectors throughout one generation. Therefore, the players that will become x_0 are those x_0 who have been selected as recipient in this round, plus x_0 who has been selected as donor but decided not to cooperate according to the strategy, plus x_1 who has been selected as donor but decided not to cooperate. (4.6) shows that the $\frac{x_0}{2}$ players that are recipients will stay as x_0 in the next round when meeting any type of players because recipients do not get to make choices whether to cooperate or defect. (4.7) shows that those $\frac{x_0}{2}$ players selected as donors when paired up with x_0 will stay as x_0 in the next round if, 1) the donor

sees the recipient has an image score of 0 and since the donor is a discriminator, he will defect on the recipient, and 2) the donor cannot see the recipient's image score but still decides not to cooperate with the recipient. Therefore, the possibility that the donor $\frac{x_0}{2}$ will stay as x_0 in the next round when paired up with x_0 as recipient is $q + (1-q)(1-p)$. When paired up with x_1, the $\frac{x_0}{2}$ donor will stay as x_0 in the next round if, 1) the donor sees the recipient has an image score of 1 but decides not to trust this positive image score, he will defect on the recipient and 2) the donor decides not to cooperate with the recipient when he does not see the recipient's image score. This adds up to $qy + (1-q)(1-p)$. When paired up with y_0, the donor $\frac{x_0}{2}$ cannot see the recipient's image score and so will stay x_0 in the next round if he decides not to cooperate with the recipient. This gives us $1-p$. Finally, when paired up with y_1, donors $\frac{x_0}{2}$ won't cooperate if, 1) they see that the y_1 recipient has an image score of 1 but they don't believe it, 2) image score of the recipient is unknown but the donor decides not to cooperate with a stranger. And we get $qy + (1-q)(1-p)$. In (4.8), it shows how x_1 in this round can become x_0 in the next round. To become a x_0 in the next round, a x_1 is first selected as a donor. And a x_1 donor behaves just like a x_0 donor since they both are discriminators.

Therefore we have population frequency changes like the following:

$$\frac{x_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{ will stay } x_0 \text{ next round} \tag{4.6}$$

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 \to qy + (1-q)(1-p) \\ y_0 \to 1-p \\ y_1 \to qy + (1-q)(1-p) \end{cases} \Rightarrow \text{ will stay } x_0 \text{ next round} \tag{4.7}$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 \to qy + (1-q)(1-p) \\ y_0 \to 1-p \\ y_1 \to qy + (1-q)(1-p) \end{cases} \Rightarrow \text{ will become } x_0 \text{ next round} \tag{4.8}$$

Adding up above scenarios, we get:

$$\begin{aligned} x_0(k+1) =& \frac{x_0(k)}{2} + \frac{x_0^2(k)q}{2} + \frac{x}{2}\left[(1-q)(x+y_1(k)) + y_0(k)\right](1-p) \\ &+ \frac{x_1(k)x_0(k)q}{2} + \frac{x(x_1(k)+y_1(k))qy}{2} \end{aligned} \tag{4.9}$$

And since x remains the same throughout the whole generation:

$$x_1(k+1) = x - x_0(k+1) \tag{4.10}$$

Since y_1 cooperate with no one, the only players that will become y_1 are those y_1 who have been selected as recipients and do not need to do anything in the current round:

$$\frac{y_1}{2} \text{ as recipient} \Rightarrow y_1 \text{ will stay next round} \tag{4.11}$$

Therefore we have:

$$y_1(k+1) = \frac{y_1(k)}{2} \tag{4.12}$$

and since y remains the same throughout the whole generation:

$$y_0(k+1) = y - y_1(k+1) \tag{4.13}$$

Solving above four difference equations yields population of each type in k^{th} round:

$$x_0(k) = x - \frac{x}{2}(\frac{1+x^2q}{2})^k - \frac{1-(\frac{1+x^2q}{2})^k}{1-x^2q}xp(1-xq) + \frac{1-(\frac{1+x^2q}{4})^k}{3-x^2q}(\frac{y}{2}-1+p)qxy \tag{4.14}$$

$$x_1(k) = x - x_0(k) \tag{4.15}$$

$$y_1(k) = y(\frac{1}{2})^{k+1} \tag{4.16}$$

$$y_0(k) = y - y_1(k) \tag{4.17}$$

Payoffs of all players

After knowing how population changes, we need to calculate the payoffs of each type of player. In each round, the payoff of each type of players are denoted as $P(x_0)$, $P(x_1)$, $P(y_0)$ and $P(y_1)$ and described below:

The only way to get payoff is to cooperate when selected as the donor and get $-c$ or the donor decides to cooperate with you when selected as a recipient and get b. When auditing the payoff of a player, simply look for, 1) when he will cooperate as a donor, and 2) when other players

will cooperate with him. (4.18) includes all the possible senarios that x_0 will cooperate with the recipient. When paired up with a x_0 recipient, a x_0 donor will cooperate when he cannot see the image score of the recipient and thus decides to cooperate with the stranger. When paired up with a x_1 recipient, the x_0 donor will cooperate if, 1) he sees that the recipient has an image score of 1 and also believes it, 2) he is willing to cooperate with the recipient even when the recipient's image score is unknown. When paired up with a y_0 recipient, since the image score of y_0 is always unknown, the x_0 donor will only cooperate if he decides to cooperate with a stranger. When paired up with a y_1 recipient, the x_0 donor cooperates same as that with a x_1 recipient. (4.19) shows when a x_0 is selected as the recipient, other discriminators would cooperate with him. But whitewashers, y_0 and y_1, would not cooperate with him. And since the x_0 recipient has a 0 image score, discriminators will only cooperate when they cannot see his image score and want to cooperate with a stranger. Payoffs of other players can be explained in the same way.

$$x_0 \text{ as donor when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to q(1-y)+(1-q)p \\ y_0 \to p \\ y_1 \to q(1-y)+(1-q)p \end{cases} \Rightarrow \text{will get } -c \tag{4.18}$$

$$x_0 \text{ as recipient when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to (1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{4.19}$$

$$P(x_0) = \frac{1}{2}(-c)\left[q(1-y)(x_1+y_1)+((1-q)(x+y_1)+y_0)p\right] + \frac{1}{2}bx(1-q)p \tag{4.20}$$

$$x_1 \text{ as donor when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to q(1-y)+(1-q)p \\ y_0 \to p \\ y_1 \to q(1-y)+(1-q)p \end{cases} \Rightarrow \text{will get } -c \tag{4.21}$$

$$x_1 \text{ as recipient when meeting } \begin{cases} x_0 \to q(1-y)+(1-q)p \\ x_1 \to q(1-y)+(1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{4.22}$$

$$P(x_1) = \frac{1}{2}(-c)\left[q(1-y)(x_1+y_1)+((1-q)(x+y_1)+y_0)p\right] + \frac{1}{2}bx\left[(1-y)q+(1-q)p\right] \tag{4.23}$$

$$y_0 \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{4.24}$$

$$y_0 \text{ as recipient when meeting } \begin{cases} x_0 \to p \\ x_1 \to p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{4.25}$$

$$P(y_0) = \frac{1}{2}bxp \tag{4.26}$$

$$y_1 \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{4.27}$$

$$y_1 \text{ as recipient when meeting } \begin{cases} x_0 \to q(1-y)+(1-q)p \\ x_1 \to q(1-y)+(1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{4.28}$$

$$P(y_1) = \frac{1}{2}bx\left[q(1-y)+(1-q)p\right] \tag{4.29}$$

Payoffs of discriminators and whitewashers

Finally we can compute discriminator's and defector's payoff in the k^{th} round:

$$\begin{aligned} P_{disc}(k) =& \frac{x_0}{x}p(x_0) + \frac{x_1}{x}p(x_1) \\ =& \left[\frac{1}{2}(\frac{1+x^2q}{2})^k + \frac{1-(\frac{1+x^2q}{2})^k}{1-x^2q}p(1-xq) - \frac{1-(\frac{1+x^2q}{4})^k}{3-x^2q}qxy(\frac{y}{2}-1+p)\right](\frac{b-c}{2})x^2q \\ &+ \frac{1}{2}cp(1-xq) + \frac{1}{2}bx(1-q)p - \frac{1}{2}cqy_1(x-p) \end{aligned} \tag{4.30}$$

$$P_{ww}(k) = \frac{1}{2}bx\left[q(1-y)+(1-q)p\right] + (\frac{1}{2})^{k+1}\frac{1}{2}bxq(q-x) \tag{4.31}$$

Summing up all their payoffs in each round, we get their payoff for the whole generation:

$$\begin{aligned} P_{disc} =& \frac{1}{1-w}\left\{\frac{1}{2}cp(1-xq) + \frac{1}{2}bx(1-q)p + \frac{b-c}{2}x^2q(\frac{xp(1-xq)}{1-x^2q} - \frac{qxy(\frac{y}{2}-1+p)}{3-x^2q})\right\} \\ &+ \frac{1}{4(2-w)}cq(x-p)(1-x) \\ &+ \frac{b-c}{2}x^2q\left\{\frac{1+x^2q}{4-2w(1+x^2q)} - \frac{p(1-xq)}{1-x^2q}\frac{1+x^2q}{2-w(1+x^2q)} + \frac{qxy(\frac{y}{2}-1+p)}{3-x^2q}\frac{1+x^2q}{4-w(1+x^2q)}\right\} \end{aligned} \tag{4.32}$$

$$P_{ww} = \frac{bx}{2(1-w)}\left[qx+(1-q)p\right] + \frac{bxq(p-x)}{4-2w} \tag{4.33}$$

4.2.3 Social identity

With social identity, a player can be paired up with only a direct friend or "a friend of a friend's", which means either you have been paired up with this player before or you've been paired up once with one of his direct friend. And once two players have been paired up once, we connect them as direct friends. Unlike pseudonym with which image score can be forged, using social identities, you can trust the image score you see about your opponents, because it is not your opponent that's telling you about himself. It is other players on the network that are telling you about your opponent. Therefore, the opponent's image score a player sees in a social identity system is more trustworthy, thus, discriminators won't have to decide whether to trust an image score of 1 or not.

Population changes

Like in section 4.2.2, we first model the population changes with social identity. Note that, y_1's image score should be visible to others like discriminators'. The $\frac{x_0}{2}$ recipient will stay as x_0 in the

next round. Therefore, the possibility of x_0 donor to stay as x_0 in the next round when paired up with any player is 1, as shown in (4.34). When x_0 is selected as the donor, he will defect on a x_0 recipient if, 1) he sees that x_0 has an image score of 0, 2) he cannot see the recipient's image score but decides to defect on the stranger. When paired up with a x_1 recipient, he will cooperate unless he does not know the recipient's image score and decide to defect on the stranger. Since y_0's image score is always unknown, the x_0 donor will defect on a y_0 recipient when he does not want to cooperate with the stranger. As y_1 has the same image score with x_1, a y_1 recipient will get the same treatment from a x_0 donor. (4.35) summaries all the situations a x_0 donor will stay as x_0 in the next round. In (4.36), it is shown how x_1 is becoming x_0. When x_0 is selected as the donor, the donor will defect on a x_0 recipient if, 1) he sees x_0 has an image score of 0, 2) he doesn't know the recipient's image score and still decides to defect. When paired up with a x_1 recipient, the donor will only defect if the recipient's image score is unknown and the donor does not want to cooperate with the stranger. When paired up with y_0, the donor will defect if he doesn't want to cooperate with a stranger. When paired up with a y_1 recipient, the donor cooperates with the same probability as with a x_1 recipient.

$$\frac{x_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.34}$$

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 \to (1-q)(1-p) \\ y_0 \to 1-p \\ y_1 \to (1-q)(1-p)) \end{cases} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.35}$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 \to (1-q)(1-p) \\ y_0 \to 1-p \\ y_1 \to (1-q)(1-p)) \end{cases} \Rightarrow \text{will become } x_0 \text{ next round} \tag{4.36}$$

Adding up all the possible scenarios:

$$x_0(k+1) = \frac{x_0(k)}{2} + \frac{x_0^2(k)q}{2} + \frac{x}{2}\left[(1-q)(x+y_1(k)) + y_0\right](1-p) + \frac{x_1(k)x_0(k)q}{2} \tag{4.37}$$

$$x_1(k+1) = x - x_0(k+1) \tag{4.38}$$

$$y_1(k+1) = \frac{y_1(k)}{2} \tag{4.39}$$

$$y_0(k+1) = y - y_1(k+1) \tag{4.40}$$

Solving the difference equations above yields:

$$x_0(k) = x\left[(\frac{1+xq}{2})^k(p-\frac{1}{2}) + (1-p)\right] - \frac{1-(\frac{1+xq}{4})^k}{3-xq}x(1-x)(1-p)q \tag{4.41}$$

$$x_1(k) = x - x_0(k) \tag{4.42}$$

$$y_1(k) = y(\frac{1}{2})^{k+1} \tag{4.43}$$

$$y_0(k) = y - y_1(k) \tag{4.44}$$

Payoffs of all players

In each round, each type of player's payoff is as follows. Taking x_0 as an example, when selected as a donor, a x_0 donor will cooperate with a x_0 recipient when he cannot see the image score of the recipient and still wants to cooperate with a stranger. With a x_1 recipient, the x_0 donor will cooperate if, 1) the image score of the recipient is unknown but the donor still decides to cooperate, 2) the donor sees that a x_1 recipient has an image score of 1. When paired up with a y_0 recipient, the only way that the donor will cooperate is when the donor wants to cooperate with a stranger. When paired up with a y_1, the x_0 donor cooperates as with a x_1 recipient.

$$x_0 \text{ as donor when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to (1-q)p + q \\ y_0 \to p \\ y_1 \to (1-q)p + q \end{cases} \Rightarrow \text{will get } -c \tag{4.45}$$

$$x_0 \text{ as recipient when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 \to (1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{4.46}$$

$$P(x_0) = \frac{1}{2}(-c)\left[q(x_1+y_1) + p(1-q)(x+y_1) + py_0\right] + \frac{1}{2}bx(1-q)p \tag{4.47}$$

$$x_1 \text{ as donor when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 \to (1-q)p + q \\ y_0 \to p \\ y_1 \to (1-q)p + q \end{cases} \Rightarrow \text{will get } -c \tag{4.48}$$

$$x_1 \text{ as donor when meeting} \begin{cases} x_0 \to (1-q)p + q \\ x_1 \to (1-q)p + q \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{4.49}$$

$$P(x_1) = \frac{1}{2}(-c)\left[q(x_1+y_1) + p(1-q)(x+y_1) + py_0\right] + \frac{1}{2}bx\left[q + (1-q)p\right] \tag{4.50}$$

$$y_0 \text{ as donor when meeting} \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{4.51}$$

$$y_0 \text{ as recipient when meeting} \begin{cases} x_0 \to p \\ x_1 \to p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{4.52}$$

$$P(y_0) = \frac{1}{2}bxp \tag{4.53}$$

$$y_1 \text{ as donor when meeting} \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{4.54}$$

$$y_1 \text{ as recipient when meeting} \begin{cases} x_0 \to q+(1-q)p \\ x_1 \to q+(1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{4.55}$$

$$P(y_1) = \frac{1}{2}bx\left[q+(1-q)p\right] \tag{4.56}$$

Payoffs of discriminators and whitewashers

Each type of player's payoff in the k^{th} round:

$$\begin{aligned} P_{disc}(k) =& \frac{x_0}{x}p(x_0) + \frac{x_1}{x}p(x_1) \\ =& -\frac{1}{2}bxq\left[(\frac{1+xq}{2})^k(p-\frac{1}{2}-\frac{1-(\frac{1+xq}{4})^k}{3-xq}q(1-x)(1-p)\right] + \frac{1}{2}px(b-c) \\ & -\frac{1}{2}c\left[q(x_1+y_1)-pqx+py-pqy_1\right] \end{aligned} \tag{4.57}$$

$$P_{ww}(k) = \frac{1}{2}bxp + \frac{1}{2}(\frac{1}{2})^{k+1}bxq(1-p) \tag{4.58}$$

Summing up all their payoffs in each round, we get their payoff for the whole generation:

$$\begin{aligned} P_{disc} =& bxq(\frac{1}{2}-p)\frac{1}{2-w(1+xq)} - \frac{(1-x)(1-p)q}{3-xq}\frac{2bxq}{4-w(1+xq)} \\ & + \frac{b+c}{2}\frac{xq}{1-w}\frac{(1-x)(1-p)q}{3-xq} + \frac{bxp}{2(1-w)} \\ & - \frac{c}{2}\left[\frac{p-2pqx}{1-w} + \frac{qx}{1-w} + \frac{2(p-\frac{1}{2})qx}{2-w(1+xq)} + \frac{4xq^2(1-x)(1-p)}{(3-xq)(4-w(1+xq))} + \frac{q(1-p)(1-x)}{2-w}\right] \end{aligned} \tag{4.59}$$

$$P_{ww} = \frac{bxp}{2(1-w)} + \frac{bx(1-p)q}{8-4w} \tag{4.60}$$

4.3 Pay your due

With PYD, upon joining in the game, a player has to pay dues to other veteran players. Dues are paid by cooperating with the recipient and each time one pays c. We require dues to be paid up to b, that is, a newcomer has to keep paying dues for $\frac{b}{c}$ rounds. Note that dues are paid to only veterans. When two newcomers are paired up, no one pays due to any other. After paying all his dues, a player becomes a veteran. The veteran cooperates with each other when paired up. We use a label of 1 and 0 to distinguish veterans and players who are still paying their dues. Discriminators are players that follow the strategy above. They pay dues only to veterans and after they become veterans, they keep cooperating with other veterans. y_1 never cooperate with others when selected as donor, and change identity soon after the defection as a donor. So donors in y_1 will become y_0 in next round. Whereas y_0 pays dues to everyone no matter if the recipient has flag of 1 or 0 because once y_0 have paid all their dues, they become y_1 and can never cooperate. They can just wait for others to pay dues to them.

4.3.1 Fixed identity

Like in section 4.2.1, we set $p = 0$, $q = 1$. The population frequency changes are as follow. (4.61) shows when a x_0 player is selected as the donor, he gets to pay his dues and if he is paired up with another x_0 player, he does not pay dues to the recipient. If paired up with a x_1 player, the probability that x_0 will stay as x_0 in the next round equals the probability that he hasn't paid all his dues after this time, which is $1 - \frac{c}{b}$. When paired up with a y_0 recipient, x_0 does not pay dues and will stay as a x_0 player in the next round. If paired up with a y_1 recipient, he stays as a x_0 player in the next round with a probability of $1 - \frac{c}{b}$. If the x_0 player is selected as a recipient (4.62), he will stay as a x_0 player in the next round. There is no way a x_1 can become a x_0.

$$\frac{x_0}{2} \text{ PYD when meeting} \begin{cases} x_0 \to 0 \\ x_1 \to 1 - \frac{c}{b} \\ y_0 \to 0 \\ y_1 \to 1 - \frac{c}{b} \end{cases} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.61}$$

$$\frac{x_0}{2} \text{ as recipient} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.62}$$

"PYD" means this player has not paid all his dues yet and is still in the due-paying process. Adding up all the possible scenarios:

$$x_0(k+1) = \frac{1}{2}x_0(k) + \frac{1}{2}x_0(k)(x_1(k) + y_1(k))(1 - \frac{c}{b}) \tag{4.63}$$

As for y_0, they would want to pay their dues and become a veteran as fast as possible so that they can defect on others. Therefore, we have the population changes of y_0 below:

$$\frac{y_0}{2} \text{ PYD when meeting } \begin{cases} x_0 \to 1-\frac{c}{b} \\ x_1 \to 1-\frac{c}{b} \\ y_0 \to 1-\frac{c}{b} \\ y_1 \to 1-\frac{c}{b} \end{cases} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.64}$$

$$\frac{y_0}{2} \text{ as recipient} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.65}$$

Adding up all the possible scenarios:

$$y_0(k+1) = \frac{1}{2}y_0(k) + \frac{1}{2}y_0(k)(1-\frac{c}{b}) \tag{4.66}$$

Now we know the population of each type of player based on its population in the last round:

$$x_0(k+1) = \frac{1}{2}x_0(k) + \frac{1}{2}x_0(k)(x_1(k)+y_1(k))(1-\frac{c}{b}) \tag{4.67}$$

$$x_1(k+1) = x - x_0(k+1) \tag{4.68}$$

$$y_0(k+1) = \frac{1}{2}y_0(k) + \frac{1}{2}y_0(k)(1-\frac{c}{b}) \tag{4.69}$$

$$y_1(k+1) = y - y_0(k+1) \tag{4.70}$$

We include the payoffs of each type of players in Appendix A.1. Unlike with image score, the difference equations we got after combining population changes and payoffs are not linear. Therefore, we choose to simulate the dynamic using Matlab.

4.3.2 Pseudonym

Since only whitewashers will try to fake their due paying label, discriminators may not trust a due-payment label of 1, and this probability is y, which is the population of whitewashers. Since a x_1 player will never become a x_0 player, the only population that will turn to x_0 in the next round are those x_0 who are selected as the recipients (4.72) and those who refuse to cooperate when selected as the donor (4.71). After a x_0 player is selected as the donor, when he's paired up with a x_0 recipient, he will stay as a x_0 in the next round if, 1) he sees that the recipient is still paying dues: q, 2) he does not know whether the recipient has paid all his dues or not but he decides to cooperate with this stranger but after this payment he still needs to pay dues: $(1-q)p(1-\frac{c}{b})$, 3) he does not know the recipient's label and does not want to cooperate with a stranger: $(1-q)(1-p)$. When the

x_0 donor is paired up with a x_1 recipient, he will not cooperate if, 1) he sees that the recipient is a veteran and trusts the recipient but still hasn't paid all his dues after this payment: $q(1-y)(1-\frac{c}{b})$, 2) he does not know about the recipient's due-paying status but is willing to cooperate with the stranger and this is not his last due to pay: $(1-q)p(1-\frac{c}{b})$, 3) he does not know the recipient's label and is not willing to cooperate with a stranger: $(1-q)(1-p)$, 4) he knows that the recipient is a veteran but does not believe it: qy. When paired up with a y_0 recipient, the x_0 donor behaves just like that with a x_0 recipient. Note that in image score, y_0's image score is always unknown because they are always changing identities after they defect. But here, y_0 have to stay with the same identity before they become y_1. A y_1 recipient will get the same response from a x_0 donor as with a x_1 recipient.

$$\frac{x_0}{2} \text{ PYD when meeting } \begin{cases} x_0 \to q + (1-q)p(1-\frac{c}{b}) + (1-q)(1-p) \\ x_1 \to q(1-y)(1-\frac{c}{b}) + (1-q)p(1-\frac{c}{b}) + (1-q)(1-p) + qy \\ y_0 \to q + (1-q)p(1-\frac{c}{b}) + (1-q)(1-p) \\ y_1 \to q(1-y)(1-\frac{c}{b}) + (1-q)p(1-\frac{c}{b}) + (1-q)(1-p) + qy \end{cases} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.71}$$

$$\frac{x_0}{2} \text{ as recipient} \Rightarrow \text{will stay } x_0 \text{next round} \tag{4.72}$$

Adding up all the possible scenarios, we get:

$$\begin{aligned} x_0(k+1) =& \frac{1}{2}x_0(k) + \frac{1}{2}x_0(k)\{(x_0(k)+y_0(k))q + (1-q)p(1-\frac{c}{b}) \\ &+ (1-q)(1-p) + (x_1(k)+y_1(k))[q(1-\frac{c}{b})(1-y) + qy]\} \end{aligned} \tag{4.73}$$

For y_0, like in fixed identity, they pay dues to everyone.

$$\frac{y_0}{2} \text{ PYD when meeting } \begin{cases} x_0 \to (1-\frac{c}{b}) \\ x_1 \to (1-\frac{c}{b}) \\ y_0 \to (1-\frac{c}{b}) \\ y_1 \to (1-\frac{c}{b}) \end{cases} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.74}$$

$$\frac{y_0}{2} \text{ as recipient} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.75}$$

$$y_0(k+1) = \frac{1}{2}y_0(k) + \frac{1}{2}y_0(k)(1-\frac{c}{b}) + \frac{1}{2}y_1(k) \tag{4.76}$$

Now we know the population of each type of player based on its population in the last round:

$$\begin{aligned} x_0(k+1) =& \frac{1}{2}x_0(k) + \frac{1}{2}x_0(k)[(x_0(k)+y_0(k))q + (1-q)p(1-\frac{c}{b}) \\ &+ (1-q)(1-p) + (x_1(k)+y_1(k))(q(1-\frac{c}{b})(1-y)+qy)] \end{aligned} \tag{4.77}$$

$$x_1(k+1) = x - x_0(k+1) \tag{4.78}$$

$$y_0(k+1) = \frac{1}{2}y_0(k) + \frac{1}{2}y_0(k)(1-\frac{c}{b}) + \frac{1}{2}y_1 \tag{4.79}$$

$$y_1(k+1) = y - y_0(k+1) \tag{4.80}$$

We include the payoffs of each type of players in Appendix A.2.

4.3.3 Social identity

With social identity, discriminators will always trust a veteran player. When selected as the donor and paired up with a x_0 recipient, a x_0 donor (4.81) will not cooperate when, 1) he sees that the recipient still needs to pay dues to others: q, 2) he does not know whether the recipient is a veteran or not but would like to cooperate and this is not his last due to be paid: $(1-q)p(1-\frac{c}{b})$, 3) he does not know if the recipient is a veteran or still needs to pay dues to others but is willing to cooperate with him: $(1-q)(1-p)$. When paired up with a x_1 player, the x_0 donor will defect on the recipient if, 1) he knows that the recipient is a veteran and cooperates but this is not his last due to be paid: $q(1-\frac{c}{b})$, 2) he does not know whether the recipient is a veteran or not but would like to cooperate yet this is not his last due to be paid: $(1-q)p(1-\frac{c}{b})$, 3) he does not know if the recipient is a veteran and is not willing to cooperate with a stranger: $(1-q)(1-p)$, 4) he knows that the recipient is a veteran but does not believe it: qy. When paired up with a y_0 recipient, the x_0 donor defects with the same probability as with a x_0 recipient and when paired up with a y_1 recipient, the x_0 donor behaves just like with a x_1 recipient. If x_0 is selected as recipient , they will stay as x_0 in the next round (4.82).

$$\frac{x_0}{2} \text{ PYD when meeting} \begin{cases} x_0 \to q + (1-q)p(1-\frac{c}{b}) + (1-q)(1-p) \\ x_1 \to q(1-\frac{c}{b}) + (1-q)p(1-\frac{c}{b}) + (1-q)(1-p) + qy \\ y_0 \to q + (1-q)p(1-\frac{c}{b}) + (1-q)(1-p) \\ y_1 \to q(1-\frac{c}{b}) + (1-q)p(1-\frac{c}{b}) + (1-q)(1-p) + qy \end{cases} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.81}$$

$$\frac{x_0}{2} \text{ as recipient} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.82}$$

Adding up all the possible scenarios, we get:

$$\begin{aligned} x_0(k+1) =& \frac{1}{2}x_0(k) + \frac{1}{2}x_0(k)[(x_0(k)+y_0(k))q + (1-q)p(1-\frac{c}{b}) \\ &+ (1-q)(1-p) + (x_1(k)+y_1(k))q(1-\frac{c}{b})] \end{aligned} \tag{4.83}$$

For y_0,

$$\frac{y_0}{2} \text{ PYD when meeting } \begin{cases} x_0 \to (1-\frac{c}{b}) \\ x_1 \to (1-\frac{c}{b}) \\ y_0 \to (1-\frac{c}{b}) \\ y_1 \to (1-\frac{c}{b}) \end{cases} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.84}$$

$$\frac{y_0}{2} \text{ as recipient} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.85}$$

$$y_0(k+1) = \frac{1}{2}y_0(k) + \frac{1}{2}y_0(k)(1-\frac{c}{b}) + \frac{1}{2}y_1(k) \tag{4.86}$$

Now we know the population of each type of player based on its population in the last round:

$$\begin{aligned} x_0(k+1) =& \frac{1}{2}x_0(k) + \frac{1}{2}x_0(k)[(x_0(k)+y_0(k))q + (1-q)p(1-\frac{c}{b}) \\ &+ (1-q)(1-p) + (x_1(k)+y_1(k))q(1-\frac{c}{b})] \end{aligned} \tag{4.87}$$

$$x_1(k+1) = x - x_0(k+1) \tag{4.88}$$

$$y_0(k+1) = \frac{1}{2}y_0(k) + \frac{1}{2}y_0(k)(1-\frac{c}{b}) + \frac{1}{2}y_1(k) \tag{4.89}$$

$$y_1(k+1) = y - y_0(k+1) \tag{4.90}$$

We include the payoffs of each type of player in Appendix A.3.

4.4 Pavlov

With Pavlov, donors cooperates with those recipients that chose the same move as he did in the previous round, as explained in 3.3.4. If a donor who cooperated in the last round meets a recipient who also cooperated in the last round, the donor would cooperate with the recipient. If he meets a recipient who didn't cooperate in the last round, the donor would defect on the recipient. But if the donor himself didn't cooperate in the last round, he would defect on those recipients who cooperated in the last round and cooperate with those recipients who defected in the last round. Here we use a label to indicate whether a player cooperated or defected in the last round. This flag has values 1 and 0 representing those players who cooperated in the last round and defected,

respectively. A x_0 player cooperates with those recipients who have a label of 0 and defects on those who have a label of 1. x_1 does the opposite, cooperates with those who have a flag of 1 and defects on those who have a flag of 0. y_0 and y_1 always defect on the recipient whenever selected as the donor. And since y_0 always changes its identity after each defection, there is no way a donor would know about y_0's label. At first glance the strategy looks bizarre, but in our computer simulation it turned out that it always won in an environment where mistakes were likely. In the end, it was almost always the dominating strategy in the population. Almost everyone was playing Pavlov's strategy, and it was very stable; it was much better than Tit-for-Tat.

4.4.1 Fixed identity

Remember that for fixed identity: $p = 0$, $q = 1$. Population frequency changes are as follow.x_0 donors, when paired up with a x_0 recipient or a y_0 recipient will cooperate with the recipient and become x_1 in the next round. When paired up with x_1 or y_1 recipients, a x_0 donor will not cooperate and stay as a x_0 player in the next round, as shown in (4.91). Those x_0 who are selected as recipients will stay as recipients in the next round (4.92). Also x_1 donors can also become x_0 in the next round. When they are paired up with x_0 or y_0 recipients, they will defect and will become x_0 in the next round, otherwise they will cooperate and stay as x_1 players (4.93).

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 1 \\ y_0 \to 0 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.91}$$

$$\frac{x_0}{2} \text{ as recipient} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.92}$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 0 \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will become } x_0 \text{ next round} \tag{4.93}$$

Adding up all the possible scenarios:

$$x_0(k+1) = \frac{1}{2}x_0(k) + \frac{1}{2}x_0(k)(x_1(k) + y_1(k)) + \frac{1}{2}x_1(k)(x_0(k) + y_0(k)) \tag{4.94}$$

Since whitewashers are not cooperating with anyone, y_0 whether selected as donors or recipients and y_1 when selected as donors will always become y_0 in the next round.

$$\frac{y_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will get } y_0 \text{ next round} \tag{4.95}$$

$$\frac{y_0}{2} \text{ as recipient} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.96}$$

$$\frac{y_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will become } y_0 \text{ next round} \tag{4.97}$$

Adding up all the possible scenarios:

$$y_0(k+1) = y_0(k) + \frac{1}{2}y_1(k) \tag{4.98}$$

Now we know the population of each type of player based on its population in the last round:

$$x_0(k+1) = \frac{1}{2}x_0(k) + \frac{1}{2}x_0(k)(x_1(k) + y_1(k)) + \frac{1}{2}x_1(k)(x_0(k) + y_0(k)) \tag{4.99}$$

$$x_1(k+1) = x - x_0(k+1) \tag{4.100}$$

$$y_0(k+1) = y_0(k) + \frac{1}{2}y_1(k) \tag{4.101}$$

$$y_1(k+1) = y - y_0(k+1) \tag{4.102}$$

We include the payoffs of each type of player in Appendix B.1.

4.4.2 Pseudonym

As shown in (4.103), x_0 donors, when paired up with x_0 recipients will defect if, 1) the donor knows that the recipient has the same label as him but does not believe it: qy, 2) the donor does not know about the recipient's label and is not willing to cooperate with a stranger: $(1-q)(1-p)$. When paired up with a x_1 recipient, the donor will not cooperate if, 1) he sees that the recipient has cooperated in the last round: q, 2) the donor does not know whether the recipient cooperated

or defected in the last round and is not willing to cooperate with a stranger: $(1-q)(1-p)$. When paired up with a y_0 recipient, the x_0 donor will defect if he doesn't want to cooperate with a stranger, since y_0's label of 1 or 0 is always unknown. When paired up with a y_1 recipient, the x_0 donor behaves just like with a x_1 recipient. When selected as the recipient (4.104), x_0 will always stay as x_0 in the next round. When selected as the donor, x_1 can also become x_0 in the next round, as shown in (4.105).

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to qy + (1-q)(1-p) \\ x_1 \to q + (1-q)(1-p) \\ y_0 \to 1-p \\ y_1 \to q + (1-q)(1-p) \end{cases} \Rightarrow \text{will stay} x_0 \text{ next round} \tag{4.103}$$

$$\frac{x_0}{2} \text{ as recipient} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.104}$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 \to qy + (1-q)(1-p) \\ y_0 \to 1-p \\ y_1 \to qy + (1-q)(1-p) \end{cases} \Rightarrow \text{will become } x_0 \text{ next round} \tag{4.105}$$

Adding up all the possible scenarios:

$$\begin{aligned} x_0(k+1) =& \frac{1}{2}x_0(k) + \frac{x_1^2(k) + x_0^2(k)}{2}(qy + (1-q)(1-p)) + x_1(k)x_0(k)(q + (1-q)(1-p)) \\ &+ \frac{x}{2}[(1-p)y_0(k) + (1-q)(1-p)y_1(k)] + \frac{x_1(k)y_1(k)}{2}qy \end{aligned} \tag{4.106}$$

$$\frac{y_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.107}$$

$$\frac{y_0}{2} \text{ as recipient} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.108}$$

$$\frac{y_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{ will become} y_0 \text{ next round} \tag{4.109}$$

Adding up all the possible scenarios:

$$y_0(k+1) = y_0(k) + \frac{1}{2}y_1(k) \tag{4.110}$$

Now we know the population of each type of player based on its population in the last round:

$$\begin{aligned} x_0(k+1) =& \frac{1}{2}x_0(k) + \frac{x_1^2(k) + x_0^2(k)}{2}(qy + (1-q)(1-p)) + x_1(k)x_0(k)(q + (1-q)(1-p)) \\ &+ \frac{x}{2}[(1-p)y_0(k) + (1-q)(1-p)y_1(k)] + \frac{x_1(k)y_1(k)}{2}qy \end{aligned} \tag{4.111}$$

$$x_1(k+1) = x - x_0(k+1) \tag{4.112}$$

$$y_0(k+1) = y_0(k) + \frac{1}{2}y_1(k) \tag{4.113}$$

$$y_1(k+1) = y - y_0(k+1) \tag{4.114}$$

We include the payoffs of each type of player in Appendix B.2.

4.4.3 Social identity

Unlike with pseudonym, discriminators can always trust a recipient with a label of 1. When a x_0 donor is paired up with a x_0 recipient, the donor will only defect if he does not know whether the recipient has defected or cooperated in the last round and he is not willing to cooperate with a player whose label is unknown: $(1-q)(1-p)$. When paired up with a x_1 or a y_1 recipient, the x_0 donor will defect if, 1) he sees that the recipient has cooperated in the last round: q, 2) he does not know the recipient's move in the last round and does not want to cooperate with the recipient

whose move in the last round is unknown: $(1-q)(1-p)$. When paired up with a y_0 recipient, the x_0 donor will defect if he does not want to cooperate with a recipient whose move last round is unknown: $1-p$. (4.115) summarizes all the above scenarios. And when selected as a recipient, x_0 players will stay as x_0 in the next round (4.116). Another type of player that will become x_0 in the next round are those x_1 players who are selected as donors. (4.117) explains it in more details. When paired up with a x_0 recipient, the x_1 donor will defect if, 1) he sees that the recipient has defected in the last round: q, 2) he does not know whether the recipient cooperated or not and is not willing to cooperate with a recipient whose last move is unknown: $(1-q)(1-p)$. When paired up with a x_1 or y_1 recipient, the x_1 donor will defect if he does not know whether the recipient cooperated or defected and he does not want to cooperate with a stranger: $(1-q)(1-p)$. When paired up with a y_0 recipient, the x_1 donor will defect on the recipient if he does not want to cooperate with a stranger.

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to (1-q)(1-p) \\ x_1 \to q + (1-q)(1-p) \\ y_0 \to 1-p \\ y_1 \to q + (1-q)(1-p) \end{cases} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.115}$$

$$\frac{x_0}{2} \text{ as recipient} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.116}$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 \to (1-q)(1-p) \\ y_0 \to 1-p \\ y_1 \to (1-q)(1-p) \end{cases} \Rightarrow \text{will become } x_0 \text{ next round} \tag{4.117}$$

Adding up all the possible scenarios:

$$x_0(k+1) = \frac{1}{2}x_0(k) + \frac{x}{2}(1-p)[y_0(k) + (1-q)(x + y_1(k))] + x_1(k)x_0(k)q + \frac{x_0(k)y_1(k)}{2}q \tag{4.118}$$

$$\frac{y_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.119}$$

$$\frac{y_0}{2} \text{ as recipient} \Rightarrow \text{will stay } y_0 \text{ next round} \tag{4.120}$$

$$\frac{y_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will become } y_0 \text{ next round} \tag{4.121}$$

Adding up all the possible scenarios:

$$y_0(k+1) = y_0(k) + \frac{1}{2}y_1(k) \tag{4.122}$$

Now we know the population of each type of player based on its population in the last round:

$$\begin{aligned} x_0(k+1) =& \frac{1}{2}x_0(k) + \frac{x}{2}(1-p)[y_0(k) + (1-q)(x + y_1(k))] + x_1(k)x_0(k)q \\ &+ \frac{x_0(k)y_1(k)}{2}q \end{aligned} \tag{4.123}$$

$$x_1(k+1) = x - x_0(k+1) \tag{4.124}$$

$$y_0(k+1) = y_0(k) + \frac{1}{2}y_1(k) \tag{4.125}$$

$$y_1(k+1) = y - y_0(k+1) \tag{4.126}$$

We include the payoffs of each type of player in Appendix B.3.

4.5 Tit-for-2-tat

Every player has an image score to indicate whether he has defected twice consecutively as a donor. Essentially, TF2T is just like image score but the "image" is toggled from 1 to 0 only when this player has defected twice consecutively. The donor decides whether he's cooperating with the recipient based on the image score of the recipient. Image score has values of only 1 or 0. If a player has defected twice consecutively as a donor, he will be given an image score of 0, otherwise he has 1. Donors are supposed to only cooperate with those recipients who have image scores of 1, otherwise donors do not cooperate. Thus, whitewashers can change their identities after defecting

twice. The population changes of whitewashers is different from that in the three strategies above. We analyze how the whitewasher population changes in section 4.5.1.

4.5.1 Fixed identity

Since a player has to stay with only one unique identity, a discriminator can always trust a recipient with an image score of 1. The population changes are as follow. z denotes those players who have defected once, but since TF2T only assigns a 0 image score when a player defected twice consecutively, z will still be treated like a x_1 and they still have 1 as their image score. As shown in (4.127), when paired up with a x_0 or y_0 recipient, a x_0 donor will defect on the recipient seeing that the recipient's image score is 0, which means the recipient did not cooperate in the last two times when selected as a donor. When a x_0 donor is paired up with a x_1 or z or y_1 recipient, the x_0 donor will cooperate with them because they have 1 as their image scores. When selected as recipients (4.128), x_0 will always stay as x_0 players in the next round. Besides, z can also become x_0 (4.130), exactly as x_1 players become z players. When paired up with x_0 or y_0, x_1 will never cooperate and becomes z in the next round (4.129), otherwise, x_1 will stay as x_1 in the next round. Also, z players that have been selected as recipients will stay as z in the next round as well, as shown in (4.131).

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 + z \to 0 \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.127}$$

$$\frac{x_0}{2} \text{ as recipient} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.128}$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 + z \to 0 \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will become } z \text{ next round} \tag{4.129}$$

$$\frac{z}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 + z \to 0 \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will become } x_0 \text{ next round} \tag{4.130}$$

$$\frac{z}{2} \text{ as recipient} \Rightarrow \text{will stay } z \text{ next round} \tag{4.131}$$

Now we have:

$$x_0(k+1) = \frac{x_0(k)}{2} + \frac{x_0(k)+z(k)}{2}(x_0(k)+y_0(k)) \tag{4.132}$$

$$z(k+1) = \frac{x_1(k)}{2}(x_0(k)+y_0(k)) + \frac{z(k)}{2} \tag{4.133}$$

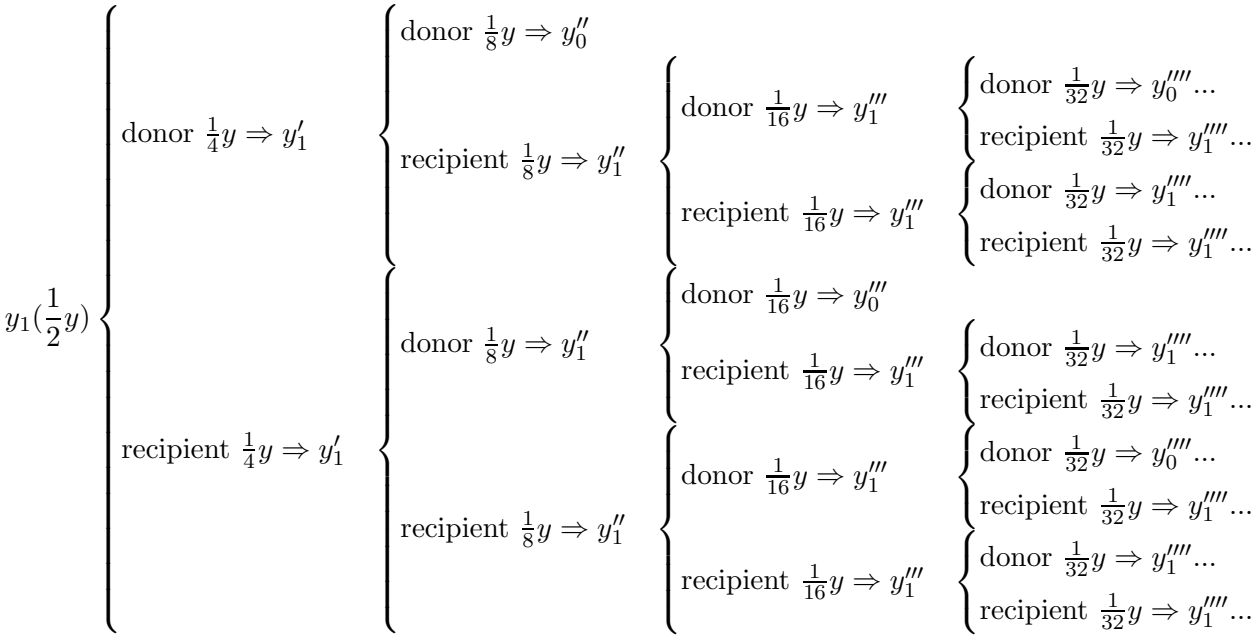
$$x_1(k+1) = x - x_0(k+1) - z(k+1) \tag{4.134}$$

As for population changes of the whitewashers, since one's image score is set to 0 after defecting two times in a row, we can predict how many y_1 will become y_0 in a certain round. Note that whitewashers never cooperate, so there is no way y_0 could become y_1. The proportion of y_1 that turns into y_0 depends on the last two rounds. Therefore we have the following table summarizing how y_1 becomes y_0 in each round:

No. of round / Whitewashers	1st	2nd	3rd	4th	5th
y1/y	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}-\frac{1}{8}$	$\frac{1}{2}-\frac{1}{8}-\frac{1}{16}$	$\frac{1}{2}-\frac{1}{8}-\frac{1}{16}-\frac{2}{32}$
y0/y	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}+\frac{1}{8}$	$\frac{1}{2}+\frac{1}{8}+\frac{1}{16}$	$\frac{1}{2}+\frac{1}{8}+\frac{1}{16}+\frac{2}{32}$

Table 4.1 Changing of Whitewashers

And the following diagram shows how y_1 becomes y_0 in detail. Note that this process holds for all three ID types.



We use two serials to simulate this change:

$$A(k) = \begin{cases} 1 & (k = 1) \\ 2 & (k = 2) \\ (A(k-1) - B(k-1)) \times 2 & (k = 3, 4, 5) \end{cases}$$

$$B(k) = \begin{cases} 0 & (k = 1, 2) \\ A(k-2) - B(k-2) - B(k-1) & (k = 3, 4, 5) \end{cases}$$

The population frequency change of y_1 is as follow:

$$y_1(k+1) = y_1(k) - \frac{B(k)}{2^k} y \tag{4.135}$$

$$y_0(k+1) = y - y_1(k+1) \tag{4.136}$$

We put the payoffs of each type of players' in Appendix C.1 and simulated the evolving of the game with Matlab.

4.5.2 Pseudonym

Like with pseudonym in the three strategies above, there is a probability of faking the image scores. Therefore, discriminator may not trust an image score of 1, and this probability is y. When selected as the donor (4.137), a x_0 player will defect on a x_0 recipient if, 1) he sees that the recipient defected in the last round: q, 2) he does not know the recipient's image score but is not willing to cooperate with a stranger: $(1-q)(1-p)$. When paired up with a x_1 or z or y_1 recipient, the x_0 donor will not cooperate if 1) he knows that the recipient has an image score of 1 but does not believe it: qy, 2) he does not know the recipient's image score and does not want to cooperate with the stranger: $(1-q)(1-p)$. When paired up with a y_0 recipient, the x_0 donor will defect if he does not want to cooperate with a stranger. When selected as the recipient, x_0 will stay as x_0 in the next round (4.138). In (4.140), we can see that donors of z become x_0 in the same way as donors of x_0 stay as x_0 and donors of x_1 become z (4.139). Finally, z as recipients will stay as z in the next round

(4.141).

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 + z \to qy + (1-q)(1-p) \\ y_0 \to (1-p) \\ y_1 \to qy + (1-q)(1-p) \end{cases} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.137}$$

$$\frac{x_0}{2} \text{ as recipient} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.138}$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 + z \to qy + (1-q)(1-p) \\ y_0 \to (1-p) \\ y_1 \to qy + (1-q)(1-p) \end{cases} \Rightarrow \text{will become } z \text{ next round} \tag{4.139}$$

$$\frac{z}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 + z \to qy + (1-q)(1-p) \\ y_0 \to (1-p) \\ y_1 \to qy + (1-q)(1-p) \end{cases} \Rightarrow \text{will become } x_0 \text{ next round} \tag{4.140}$$

$$\frac{z}{2} \text{ as recipient} \Rightarrow \text{will stay } z \text{ next round} \tag{4.141}$$

Adding up all the possible scenarios, we get:

$$\begin{aligned} x_0(k+1) =& \frac{1}{2}x_0(k) + \frac{x_0(k)+z(k)}{2}[x_0(k)(q+(1-q)(1-p)) \\ &+ (x_1(k)+z(k)+y_1(k))(qy+(1-q)(1-p)) + y_0(k)(1-p)] \end{aligned} \tag{4.142}$$

$$\begin{aligned} z(k+1) =& \frac{1}{2}x_1(k)[x_0(k)(q+(1-q)(1-p)) + (x_1(k)+z(k)+y_1(k))(qy+(1-q)(1-p)) \\ &+ y_0(k)(1-p)] + \frac{1}{2}z(k) \end{aligned} \tag{4.143}$$

$$x_1(k+1) = x - x_0(k+1) - z(k+1) \tag{4.144}$$

For y_1, we can use that from 4.5.1

$$A(k) = \begin{cases} 1 & (k=1) \\ 2 & (k=2) \\ (A(k-1) - B(k-1)) \times 2 & (k=3,4,5) \end{cases}$$

$$B(k) = \begin{cases} 0 & (k=1,2) \\ A(k-2) - B(k-2) - B(k-1) & (k=3,4,5) \end{cases}$$

$$y_1(k+1) = y_1(k) - \frac{B(k)}{2^k} y \tag{4.145}$$

$$y_0(k+1) = y - y_1(k+1) \tag{4.146}$$

We put the payoffs of each type of players in Appendix C.2 and simulated the evolving of the game with Matlab.

4.5.3 Social identity

With social identity, discriminators can always believe an image score of 1. As the donor, a x_0 player will defect on a x_0 recipient and stay as a x_0 player in the next round if, 1) the donor knows that the recipient has an image score of 0: q, 2) the donor does not know the recipient's image score but does not want to cooperate with a stranger: $(1-q)(1-p)$. When paired up with a x_1 or z or y_1 recipient, the x_0 donor will defect if he does not know the recipient's image score and is not willing to cooperate with a stranger: $(1-q)(1-p)$. When paired up with a y_0 recipient, the x_0 donor will defect if he doesn't want to cooperate with a player whose image score is unknown: $(1-p)$. Conditions above are included in (4.147). x_0 recipients will always stay as x_0 players in the next round (4.148). Another type of players that might become x_0 in the next round are z donors. Since they are discriminators like x_0, they will defect on each type of recipient just like a x_0 donor (4.150). x_1 donors might become z in the next round in the same way as x_0 donors remain as x_0

in the next round. z donors will remain as x_1 players in the next round.

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 + z \to (1-q)(1-p) \\ y_0 \to (1-p) \\ y_1 \to (1-q)(1-p) \end{cases} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.147}$$

$$\frac{x_0}{2} \text{ as recipient} \Rightarrow \text{will stay } x_0 \text{ next round} \tag{4.148}$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 + z \to (1-q)(1-p) \\ y_0 \to (1-p) \\ y_1 \to (1-q)(1-p) \end{cases} \Rightarrow \text{will become } z \text{ next round} \tag{4.149}$$

$$\frac{z}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q + (1-q)(1-p) \\ x_1 + z \to (1-q)(1-p) \\ y_0 \to (1-p) \\ y_1 \to (1-q)(1-p) \end{cases} \Rightarrow \text{will become } x_0 \text{ next round} \tag{4.150}$$

$$\frac{z}{2} \text{ as recipient} \Rightarrow \text{will stay } z \text{ next round} \tag{4.151}$$

Adding up all the possible scenarios, we get:

$$x_0(k+1) = \frac{1}{2}x_0(k) + \frac{x_0(k)+z(k)}{2}\left[x_0(k)q + (x + y_1(k))(1-q)(1-p) + y_0(k)(1-p)\right] \tag{4.152}$$

$$z(k+1) = \frac{1}{2}x_1(k)\left[x_0(k)q + (x + y_1(k))(1-q)(1-p) + y_0(k)(1-p)\right] + \frac{1}{2}z(k) \tag{4.153}$$

$$x_1(k+1) = x - x_0(k+1) - z(k+1) \tag{4.154}$$

For y_1, we can use that from section4.5.1

$$A(k) = \begin{cases} 1 & (k=1) \\ 2 & (k=2) \\ (A(k-1) - B(k-1)) \times 2 & (k = 3,4,5) \end{cases}$$

$$B(k) = \begin{cases} 0 & (k=1,2) \\ A(k-2) - B(k-2) - B(k-1) & (k = 3,4,5) \end{cases}$$

$$y_1(k+1) = y_1(k) - \frac{B(k)}{2^k} y \tag{4.155}$$

$$y_0(k+1) = y - y_1(k+1) \tag{4.156}$$

We put the payoffs of each type of player in Appendix C.3 and simulated the evolving of the game with Matlab.

Chapter 5

Matlab Results

In this chapter, we give the Matlab results for our models. Our analysis of the models is threefold. Because it is the average payoff that will decide which side get to thrive and which side is doomed to be wiped out, we focus on the payoff difference between discriminators and whitewashers in the very first round of the game. The payoff difference is $P(disc) - P(ww)$, the average payoff of discriminators minus that of whitewashers. We first analyze the effect of each parameter: x, p and q on the payoff differences for fixed identity, pseudonym and social identity. Notice that a positive payoff difference means that discriminators will eliminate whitewashers in the end. And higher the payoff difference is, sooner whitewashers will be wiped out. Then we extract the population changes of all four types of players through repeated generations for each identity type and strategy combination. This will give us a more direct view of the contradictions between different identity types. Finally, we visualize the payoff a discriminator will get from the beginning until they wipe out the whitewashers. And here we define the "cost of whitewashers" as the payoff a discriminator didn't get due to whitewashers in the population.

As for the parameters: w, x, b and c, we decide their values as follow. The discount factor w symbolize how long a player is about to stay in the game and how important he weights his future payoffs. Since we want to establish a long-term secured online environment where we want our users to trust and rely on our services for as long as possible, w should not be too low. The initial value of x is fixed to 0.5 to see if the identity type is robust and can withstand a considerable amount of whitewashers. Since fixed identity gives only one identity to each player, whitewashing using fixed identity is not possible. Therefore, we use fixed identity to fix the values of b ($c = 0.1$). We first need to find the b value that makes $P(disc)_{\text{fixed identity}} = P(ww)_{\text{fixed identity}}$ for each strategy. In our analysis, we compare only pseudonym and social identity. There are two ways we use to compare the two identity types. One way is to focus on the parameter range that generates positive payoff difference. The other is to set each of them at their best and compare the value of their payoff

differences. At all times we do not set the value of any parameter over 0.9 or below 0.1. And $0 \leq p, q \leq 1$.

There are three ways for discriminators to win over whitewashers: 1) cooperate less with whitewashers, 2) cooperate more with discriminators, 3) getting more cooperations from donors when selected as the recipient but since whitewashers never cooperate with other players, a discriminator can only hope that other discriminators will cooperate with them. Since all four strategies label players with 1 or 0 and there is no way a donor will know if the recipient is a whitewasher or a discriminator, a donor can only base his decision on the recipient's label. With pseudonym, the donor may not trust a recipient with a label 1 and thus has a lower chance of cooperating with y_1 and x_1 compared to donors with social identity. However, this difference depends on q, the probability of seeing the recipient's label. Only when the donor knows about the recipient's label, he needs to decide if he trusts the label 1 or not. If q is high, trusting the label 1 will affect the payoff difference more. Another factor, y, also affects how much pseudonym's payoff difference is different from social identity's. If y is large, discriminators will not trust a recipient who has a label of 1 with pseudonym. That is to say, if the population of whitewashers is diminishing, more discriminator will trust the recipient and the difference between pseudonym and social identity is becoming smaller. We will see in later analysis that the difference between social identity and pseudonym becomes larger with increasing q, y and b (or $\frac{b}{c}$).

5.1 Image score

The b value that satisfies $P(disc)_{\text{fixed identity}} = P(ww)_{\text{fixed identity}}$ is $b = 1.5$. Thus, the parameters used here are: $w = 0.8$, $x = 0.5$, $c = 0.1$ and $b = 1.5$.

5.1.1 Payoff differences

We find that pseudonym under image score has a higher payoff difference with a higher p and a lower q. That is to say, the discriminators are not likely to know about the image score of the recipients but have to trust and cooperate with strangers a lot to defeat whitewashers. Social identity, on the other hand, is in favor of the opposite, a lower p and a higher q, which means with social identity, discriminators will often know about the image score of the recipients and do not need to blindly trust and cooperate with strangers. Clearly, social identity's way of interaction makes more sense than that with pseudonym. And the optimum point for pseudonym is ($p = 0.9$, $q = 0.1$). For social identity it is ($p = 0.1$, $q = 0.9$). From above observations, we can predict that under the present pseudonym environment, releasing more user information does not help eliminate whitewashing, since pseudonym does not work well given a lower p and a higher q and its own manner of eliminating whitewashers are against common sense. We cannot ask our user to

blindly trust unknown users.

We also want to find the area, in terms of p and q, where pseudonym has a positive payoff difference. Setting $q = 0.1$, payoff difference of pseudonym is positive when $p \geq 0.3$. Setting $p = 0.9$, we can get the range of q for pseudonym to have positive payoff difference, which is $q \leq 0.258$. Social identity's range of p and q can be found in the same way. With $q = 0.9$, $p \leq 0.4$ generates a positive payoff difference for social identity. And setting $p = 0.1$, the range of q that yields positive payoff difference for social identity is $q \geq 0.26$. Regarding their positive payoff difference range on x, we set both of them on their optimum point and find that pseudonym always has a positive payoff difference whatever x's initial value is. Social identity can only generate positive payoff difference when $x > 0.0744$.

5.1.2 Population changes

The population changes reveals the biggest contradiction between the two identity types. Figure 5.1a and Figure 5.1b shows population changes of pseudonym and social identity over 40 generations. Note that since $w = 0.8$, each generation has $\frac{1}{1-w} = 5$ rounds. Note that we set both identity types at their optimum point.

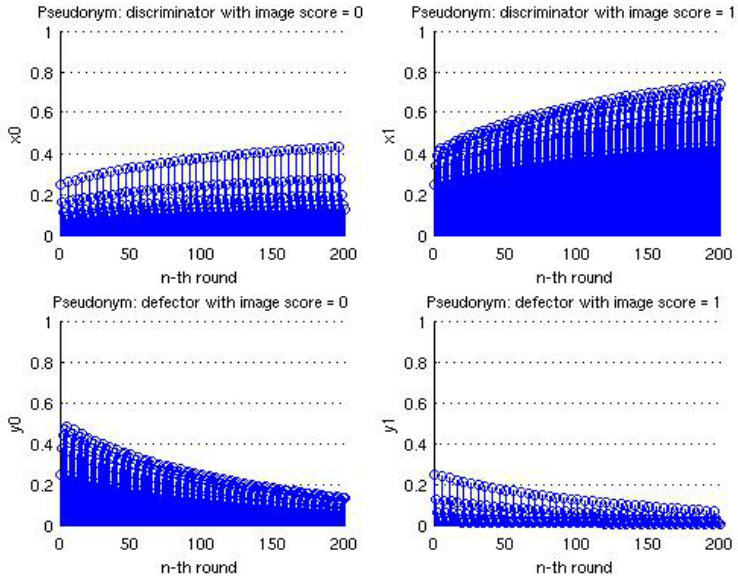


Fig. 5.1a Population Changes of Pseudonym

In both Figure 5.1a and Figure 5.1b, the four subfigures correspond to the population changes of x_0, x_1, y_0 and y_1 from top left to bottom right. At all times, four types of players population should add up to 1. We can see that with pseudonym whitewashers populations, y_0 and y_1,

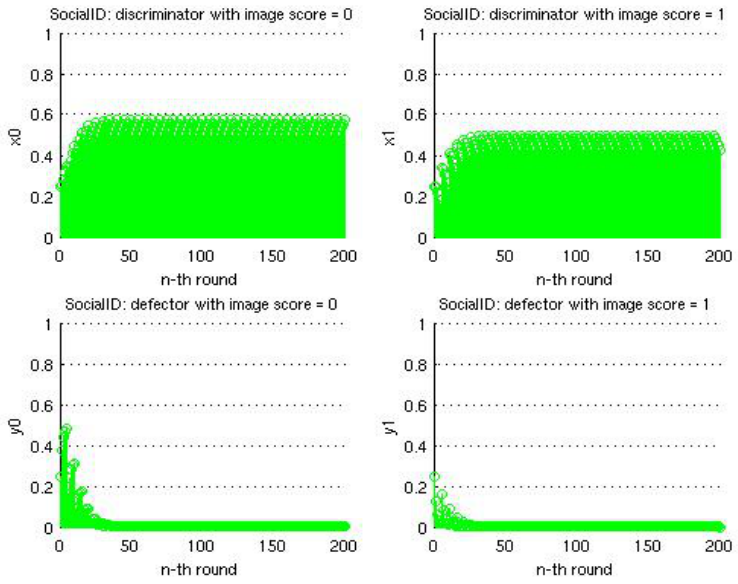


Fig. 5.1b Population Changes of Social Identity

are decreasing linearly and discriminators populations, x_0 and x_1, are increasing linearly. And through one generation the populations of x_0 and y_1 are decreasing while the populations of x_1 and y_0 are increasing, which means, discriminators have a tendency to cooperate with others and whitewashers are defecting all the time. With social identity, whitewashers are disappearing faster than in pseudonym. However, the populations of x_1 and y_1 are decreasing while the populations of x_0 and y_0 are increasing through one generation. To explain the difference of how discriminators' populations are changing through one generation, we need to refer to (4.9) and (4.37). The reason why x_0 in pseudonym is decreasing but in social identity it is increasing is that q in pseudonym is smaller than q in social identity. Remember that q is the probability that a discriminator knows about the image score of the recipient. When a discriminator as a donor is paired up with a x_0 recipient, the discriminator in pseudonym is not very likely to know about the recipient's image score of 0 and does not become a x_0 in the next round. But in social identity, the discriminator will know about the image score of the recipient and won't cooperate, thus, it becomes a x_0 in the next round. Hence, the decrease of x_1 is reasonable.

5.1.3 Cost of whitewashers

Even if one identity type can eliminate whitewashers faster than the other, we still need to see which one generates more payoff for discriminators. Figure 5.1c and Figure 5.1d show the payoff a discriminator get for each generation before the whitewashers are eliminated for pseudonym and

social identity. And we define the cost of whitewashers as the amount of payoff a discriminator could not get due to whitewashers. This amount depends on the final payoff that a discriminator gets after all whitewashers have been eliminated. We can get the cost by adding up all the payoffs a discriminator didn't get in each generation until whitewashers are wiped out.

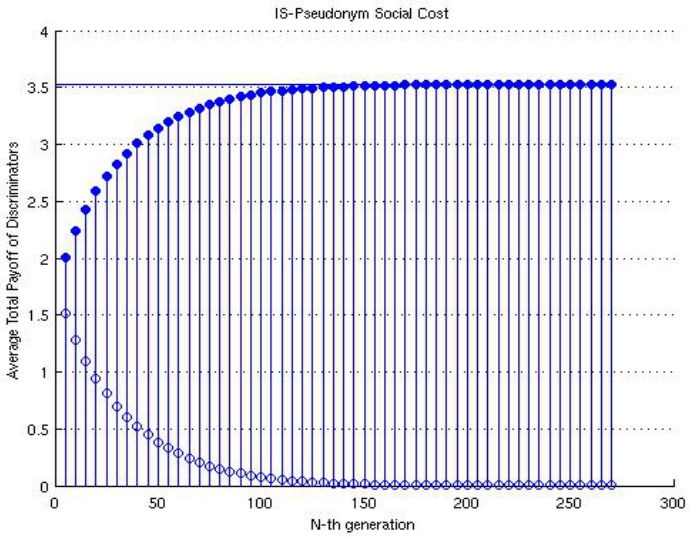


Fig. 5.1c Discriminators Payoff of Pseudonym

In Figure 5.1c and Figure 5.1d, the solid points are payoffs that a discriminator gets during a generation. The solid line is the amount of payoff a discriminator gets when there are no whitewashers anymore. And the empty points are payoffs the discriminators didn't get in each generation. From the figures we can see that pseudonym takes 272 generations to eliminate whitewashers and social identity needs only 13 generations. And pseudonym has a cost of 10.4884 to eliminate whitewashers whereas social identity costs 2.1289. However, pseudonym does have a higher final payoff for discriminators. This is because discriminators in social identity are not cooperating with themselves as much, as compared to pseudonym, which is explained in section 5.1.2.

5.2 Pay your due

We first find the value b that satisfies $P(disc)_{\text{fixed identity}} = P(ww)_{\text{fixed identity}}$ and it is $b = 0.46$. The parameters used here are: $w = 0.8$, $x = 0.5$, $c = 0.1$ and $b = 1.5$.

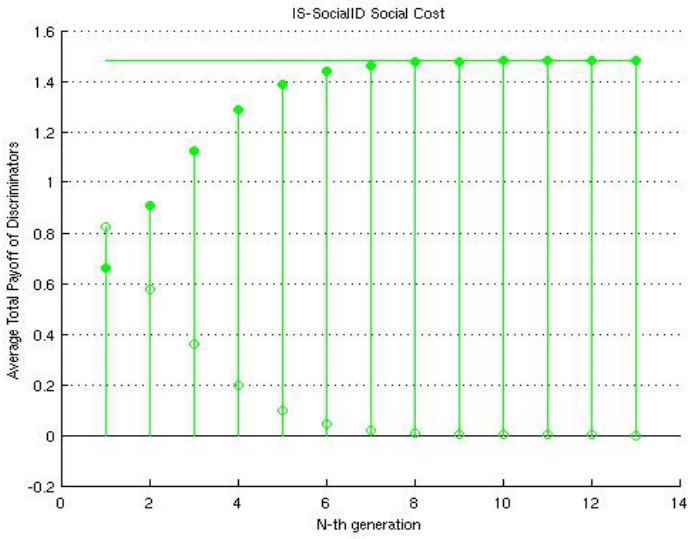


Fig. 5.1d Discriminators Payoff of Social Identity

5.2.1 Payoff differences

Unlike in image score, pseudonym and social identity have the same preference over p and q. Their payoff differences both increase with a lower p and a higher q. Hence, we use $p = 0.1$ and $q = 0.9$ as their optimum point.

To study the effect of p on the two identity types' payoff differences, we fix q to 0.9. As shown in Figure 5.2a, pseudonym and social identity both have positive payoff differences. Although both payoff differences decrease as p goes up, social identity's payoff difference decreases slower than pseudonym's. We can see that neither identity types encourage cooperating with strangers, since the payoff differences both decline with increasing p. And social identity is more tolerant about cooperating with strangers.

When extracting the effect of q on the two identity types' payoff differences, we fix p to 0.1. Figure 5.2b shows the result. Both identity types are in favor of larger q. As q increases, social identity's payoff difference grows faster than pseudonym's payoff difference. Knowing better about your recipient helps eliminating whitewashers for both social identity and pseudonym.

But unlike p which is based on users choices, q depends largely on the identity system itself. If we use pseudonym, it's not very likely that the q would be very high. And since anonymity is one of pseudonym's core advantages, it is not reasonable to ask users to release more information about themselves. In a system that uses social identity, q will be considerably higher than in pseudonym systems. A social network motivates users to share their information with those they know well and trust, and the cycle of trust is based on the connection between the users. Whenever there is

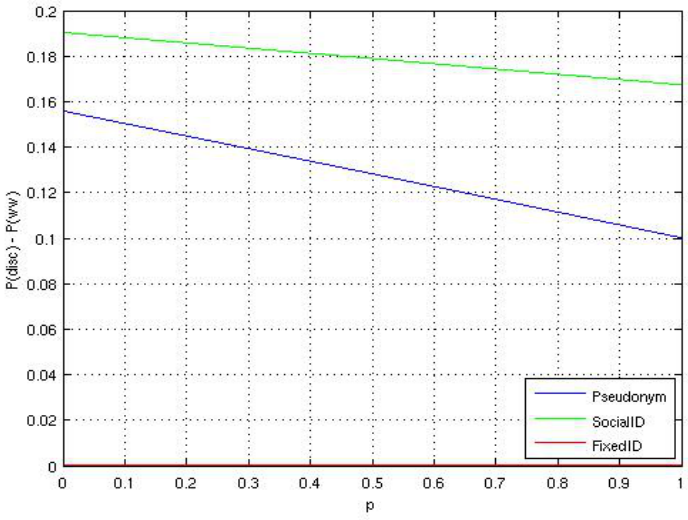


Fig. 5.2a p vs. payoff difference

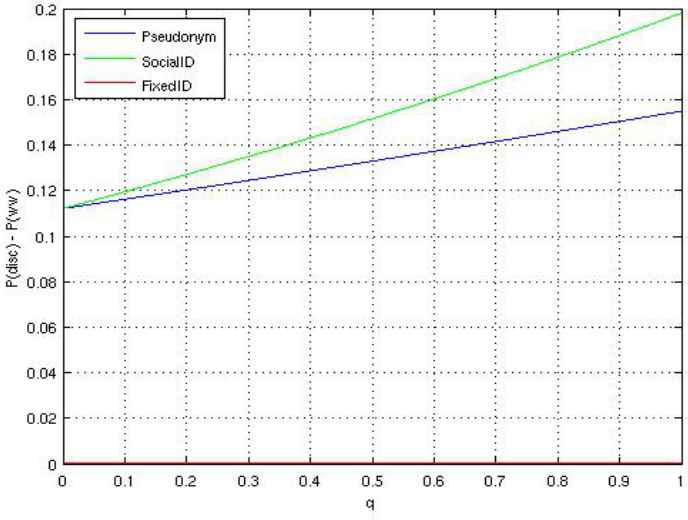


Fig. 5.2b q vs. payoff difference

a release of information, there is a good reason behind it. Therefore, users can safely manage their own information and will release it more often than in a pseudonym system.

Another interesting result is how pseudonym and social identity respond to a growing number of discriminators in the population. Figure 5.2c shows changing of payoff differences of different identity types versus x – *the initial population of discriminators* with $p = 0.1$ and $q = 0.9$.

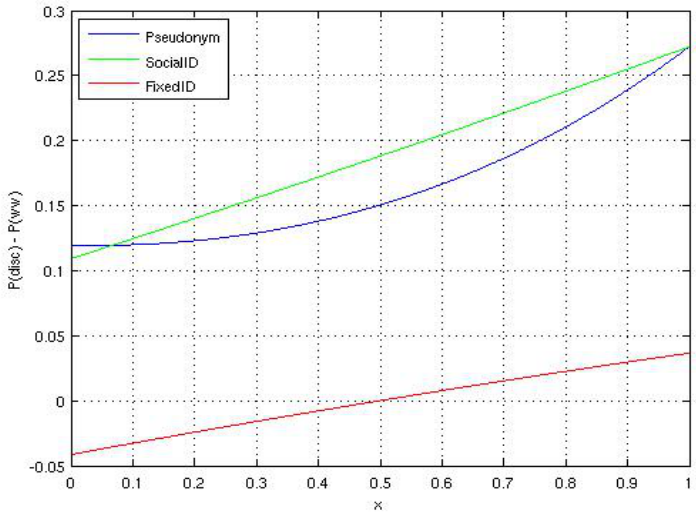


Fig. 5.2c x vs. payoff difference

Both fixed identity and social identity's payoff differences increase linearly with a growing x. The reason why pseudonym's payoff difference is lower than social identity's when $0.07 < x < 1$, is that discriminators may not believe a veteran recipient. Therefore, as whitewashers are being eliminated and the initial population of discriminators at the beginning of each generation is increasing, social identity are becoming faster at wiping out whitewashers. Notice that when there are few discriminators in the population, $x < 0.07$, pseudonym has a higher payoff difference, which is also due to discriminator not believing veteran recipients. Not trusting other people is the rightful choice when most of the users are whitewashers.

5.2.2 Population changes

Simulating different types of players population changes confirms the conclusion in section 5.2.1. Figure 5.2d and Figure 5.2e show how each type of players' population changes across 20 generations. And the parameters used are $w = 0.8$, $x = 0.5$, $c = 0.1$ and $b = 1.5$ for both pseudonym and social identity.

From Figure 5.2d and Figure 5.2e, we can see that pseudonym and social identity eliminate

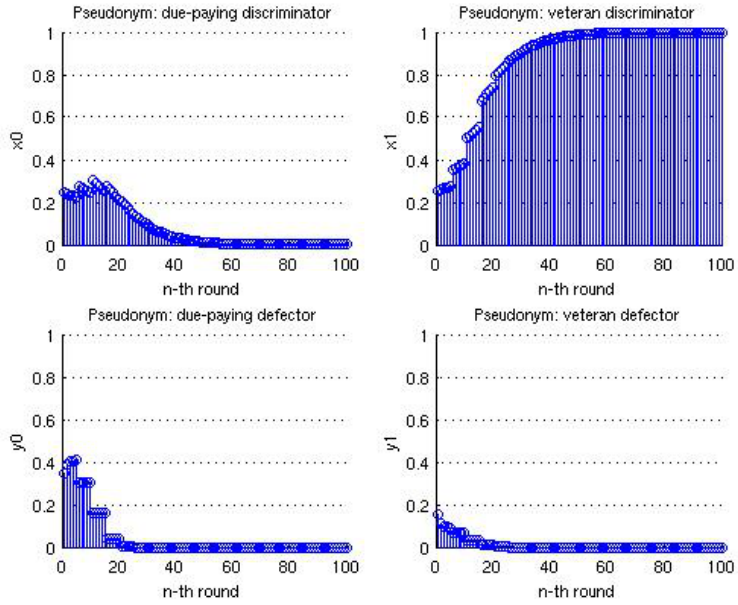


Fig. 5.2d Population Changes of Pseudonym

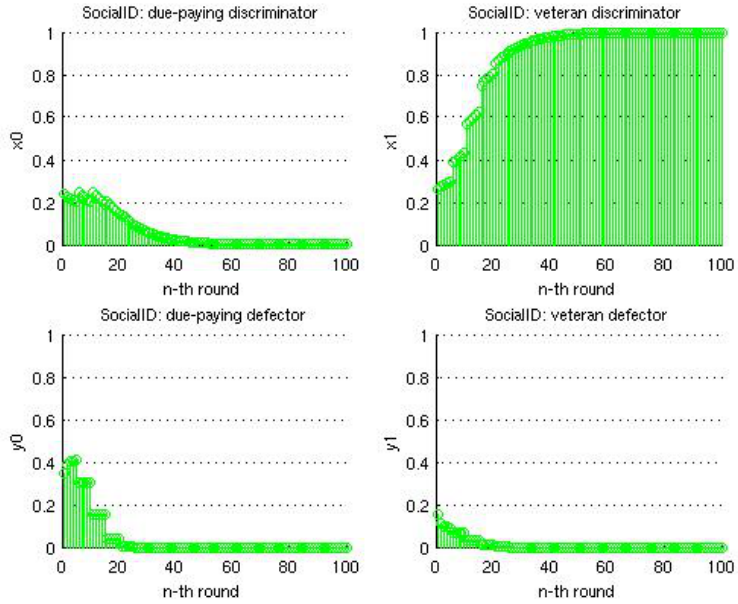


Fig. 5.2e Population Changes of Social Identity

whitewashers in the same fashion. Not only whitewashers are eliminated, all discriminators will eventually pay off all of their dues. Comparing population changes of PYD and that of image score, we can see the difference when y_0 pays its dues. Obviously, PYD is a intense strategy since everyone pays their dues first before getting anything from the community.

5.2.3 Cost of whitewashers

Calculating the cost of whitewashers for pseudonym and social identity, we find that pseudonym takes eight generations eliminating whitewashers. And social identity takes seven generations. Besides, discriminators actually get more when there are white washers in the first two generations because y_0 players are paying dues to them. But with pseudonym, discriminators are not as lucky. As shown in Figure 5.2f and 5.2g, social identity has a slightly higher final payoff for discriminators. Moreover, we sum up the payoff loss in each generation before whitewashers are eliminated. Social identity also has a lower cost of eliminating whitewashers. Cost of eliminating whitewashers in pseudonym is 1.6918 and for social identity it's 0.0189.

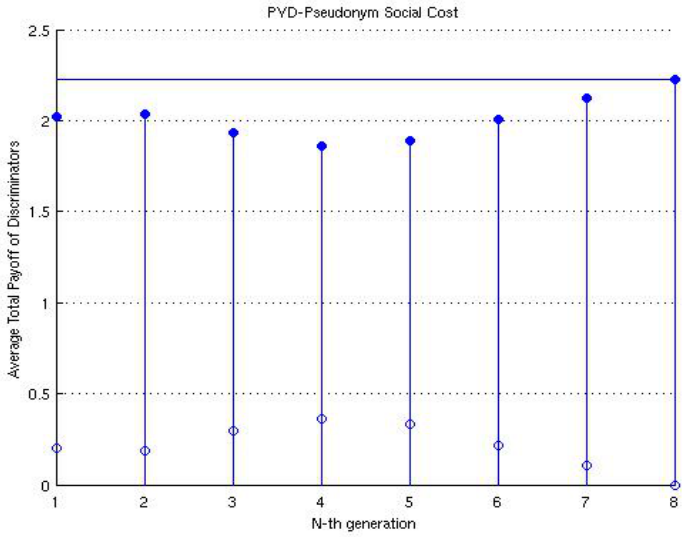


Fig. 5.2f Discriminators Payoff of Pseudonym

5.3 Pavlov

What's interesting about Pavlov strategy is that discriminators only cooperate with those who did the same thing with themselves in the last round. But defectors never cooperate and hence no loss of payoff, they wait for discriminators cooperating with them. While discriminators are

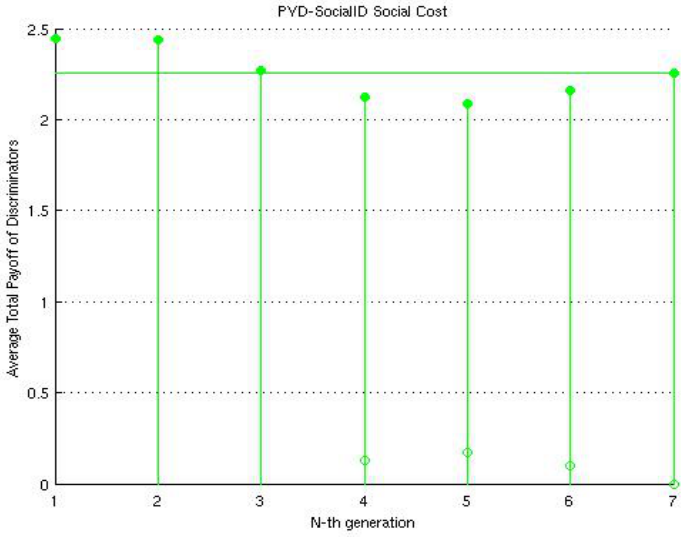


Fig. 5.2g Discriminators Payoff of Social Identity

giving payoffs to defectors, they are not getting enough to defeat defectors. Discriminators get nothing from defectors so they need to cooperate among themselves more. However, fixed identity does not encourage cooperation between discriminators. This can be also seen from the equations in Appendix B.1, $P(x_0)$ is smaller than $P(y_0)$ and $P(x_1)$ is smaller than $P(y_1)$. Therefore, fixed identity with Pavlov strategy is never evolutionary stable. We cannot find a value of b that makes payoff difference of fixed identity zero. Thus, we use $b = 1$.

5.3.1 Payoff differences

The two identity type's payoff differences both increase when p drops and q increases. We set $q = 0.9$ and find that p can be as high as 0.39 for social identity and 0.22 for pseudonym to have positive payoff difference. This means players can cooperate with strangers more with social identity than with pseudonym. Setting $p = 0.1$, we find q can be as low as 0.1 for pseudonym whereas social identity can have $q = 0.04$. Therefore, social identity can tolerate lower rate of users sharing information. Using $p = 0.1$ and $q = 0.9$, we find pseudonym has positive payoff difference when $x > 0.26$ and social identity can tolerate x to as low as 0.032. Therefore, social identity has a larger area where it yields positive payoff difference and so is more immune to whitewashers than pseudonym.

5.3.2 Population changes

Using $w = 0.8$, $p = 0.1$, $q = 0.9$, $x = 0.5$, $c = 0.1$ and $b = 1$, Figure 5.3a and Figure 5.3b are population changes of pseudonym and social identity through 20 generations. Social identity can eliminate whitewashers slightly faster than pseudonym does.

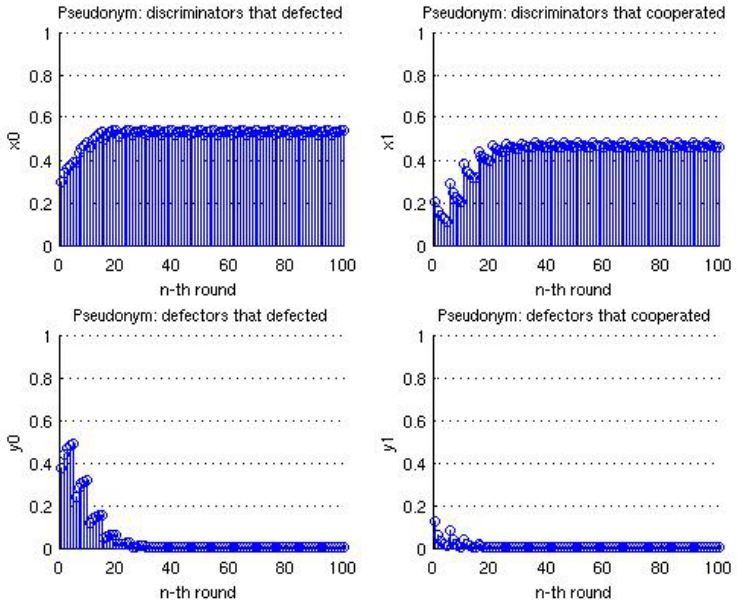


Fig. 5.3a Population Changes of Pseudonym

5.3.3 Cost of whitewashers

In Figure 5.3c, it shows pseudonym takes 10 generations to eliminate whitewashers with a cost of 1.5729 and social identity takes 10 generations and has a cost of 0.8081, as shown in Figure 5.3d. After whitewashers are eliminated, discriminators with social identity can get 1.0273 as average total payoff for each generation and discriminators with pseudonym get 1.0272 for each generation. The two identity types have the same final payoff for discriminators.

5.4 Tit-for-2-tat

Using $w = 0.8$, $x = 0.5$, $c = 0.1$, FixedID has zero payoff difference at $b = 0.4789$.

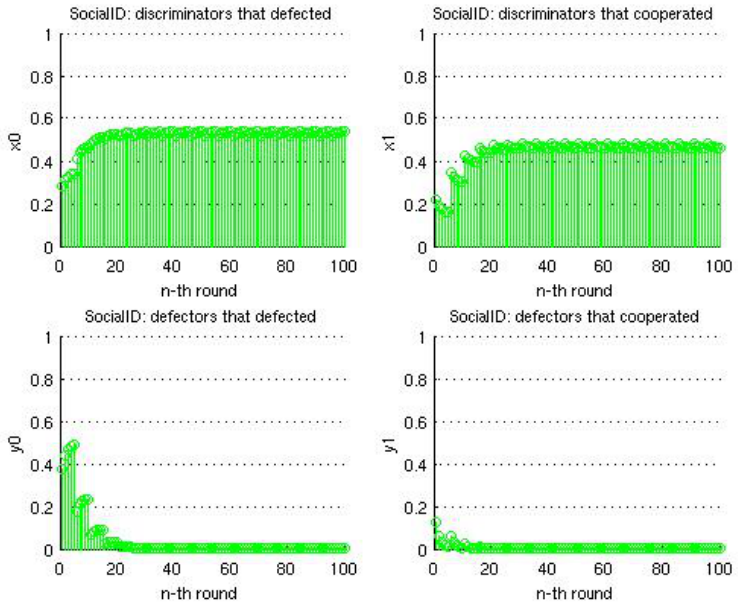


Fig. 5.3b Population Changes of Social Identity

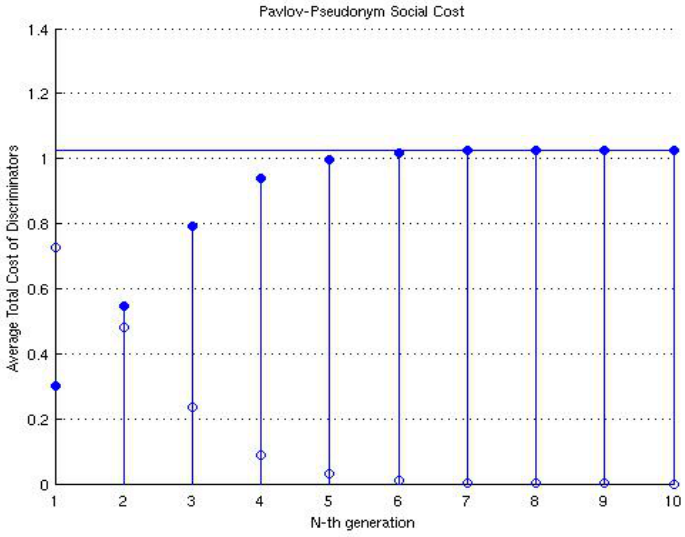


Fig. 5.3c Discriminators Payoff of Pseudonym

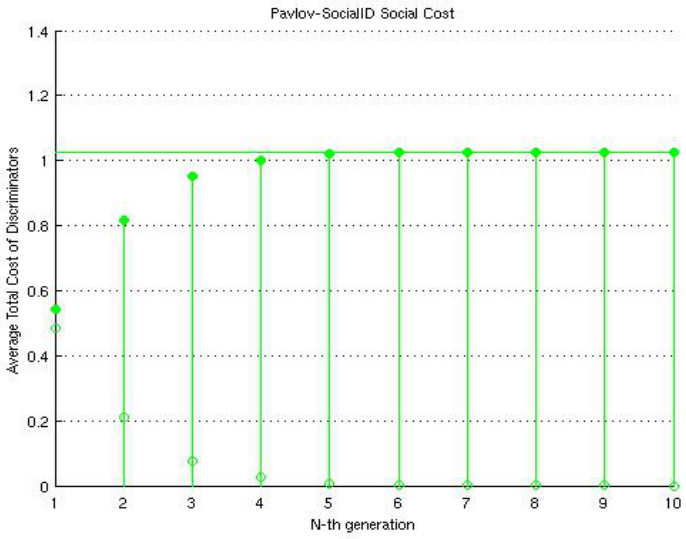


Fig. 5.3d Discriminators Payoff of Social Identity

5.4.1 Payoff differences

Using $b = 0.4789$, we find that both identity types' payoff difference increase as p decreases. However, their payoff differences are negative, as shown in Figure 5.4a. Clearly, TF2T is being too forgivable compared to image score because discriminators are defecting on recipients who have defected twice in a row. It gives whitewashers a better chance to cheat on other players.

As we can see from Figure 5.4a, payoff differences of both pseudonym and social identity are negative no matter what the value of p is and their payoff differences decrease linearly as p grows. This means that neither social identity nor pseudonym would encourage players cooperating with strangers. The reason that the two identity types' payoff differences are so close is because $q = 0.1$ and if most donors do not see the image scores of their recipients, trusting a recipient with an image score of 1 does not make much difference.

From Figure 5.4b, we can see that pseudonym's payoff difference is decreasing as q increases. Social identity's payoff difference first declines but after $q = 0.55$ it increases. If we compare the payoff of discriminators and whitewashers, as in Appendix C.2 and Appendix C.3, we can see that whitewashers always get more than discriminators. However, total payoff difference is not based on pure payoff each player gets but also population changes. The payoff difference shown in related figures is the total payoff of the first generation and during this generation if the population of either discriminators or whitewashers decrease, the total payoff they get at the end of the first generation will still decrease. The turning point, which is $q = 0.55$, makes the population of x_1 go up. And that's why social identity's payoff difference increases.

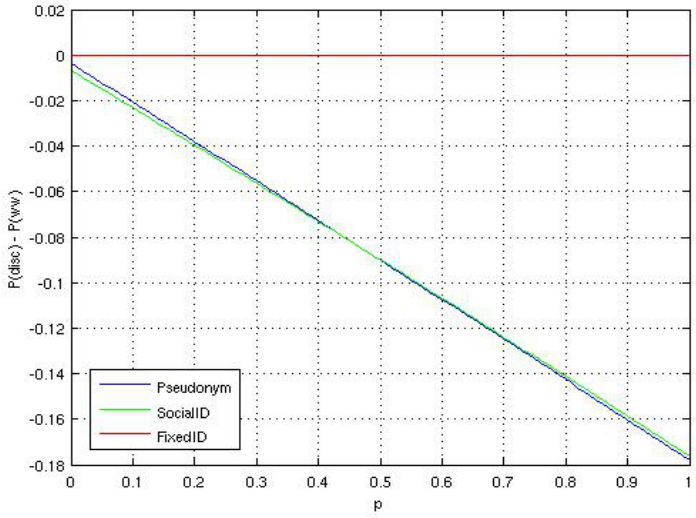


Fig. 5.4a p vs. payoff difference (b=0.4789 q=0.1)

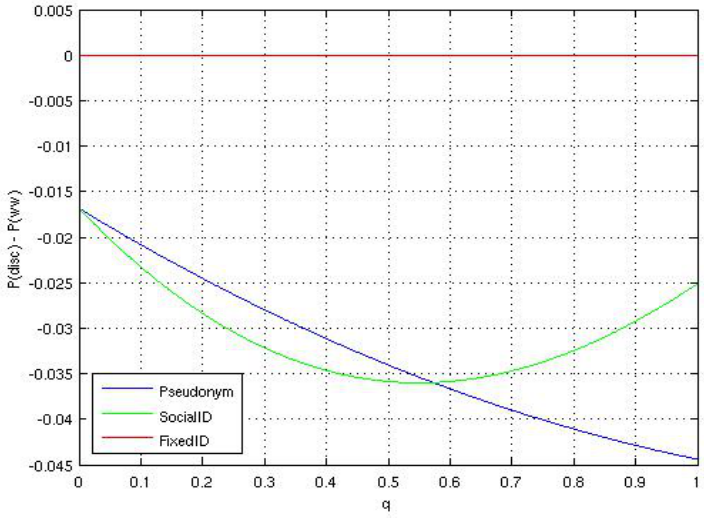


Fig. 5.4b q vs. payoff difference (b=0.4789 p=0.1)

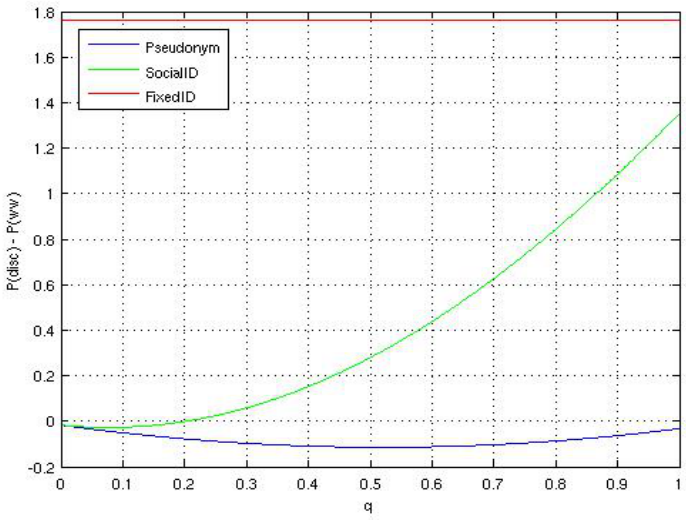


Fig. 5.4c q vs. payoff difference (b = 10)

More interestingly, if we set $b = 10$, we get Figure 5.4c. Social identity's payoff difference is positive when $q > 0.2$ and increases as q grows.

Figure 5.4d shows the relationship between x and identity types' payoff differences when $p = 0.1$ and $q = 0.1$. Fixed identity's payoff difference is 0 at $x = 0.5$ because $b = 0.684$ is the zero payoff difference point under $x = 0.5$ for fixed identity. We use $q = 0.1$ because as shown in Figure 5.4b, both identity types' payoff differences are relatively higher at $q = 0.1$. But with $p = 0.1$ and $q = 0.1$, social identity's payoff difference is hardly different from pseudonym. Since at $q = 0.9$ social identity has higher payoff difference as well, we set $q = 0.9$ and get Figure 5.4e. And Figure 5.4f shows the result after setting $b = 10$.

In both Figure 5.4e and Figure 5.4f, payoff differences of pseudonym first decrease and then increase after $x = 0.5$ as the initial population of discriminators grows. But social identity's payoff difference grows linearly with the increasing x. This is because the payoff discriminators get in social identity depends more on population changes in order to rule over whitewashers whereas in pseudonym, discriminators win over whitewashers by getting more payoff.

5.4.2 Population changes

We use $w = 0.8$, $p = 0.1$, $q = 0.9$, $x = 0.5$, $c = 0.1$ and $b = 10$ to study population changes of pseudonym and social identity. As shown in Figure 5.4g and Figure 5.4h, pseudonym cannot defeat whitewashers while social identity can help discriminators eliminate whitewashers. Moreover, if we look into x_0 and x_1 players payoff in Appendix C.2 and Appendix C.3, x_1 always gets more payoff

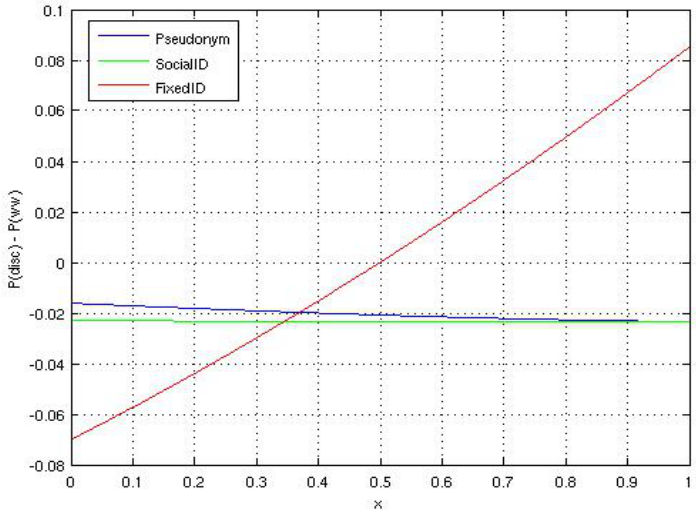


Fig. 5.4d x vs. payoff difference (p = 0.1, q = 0.1)

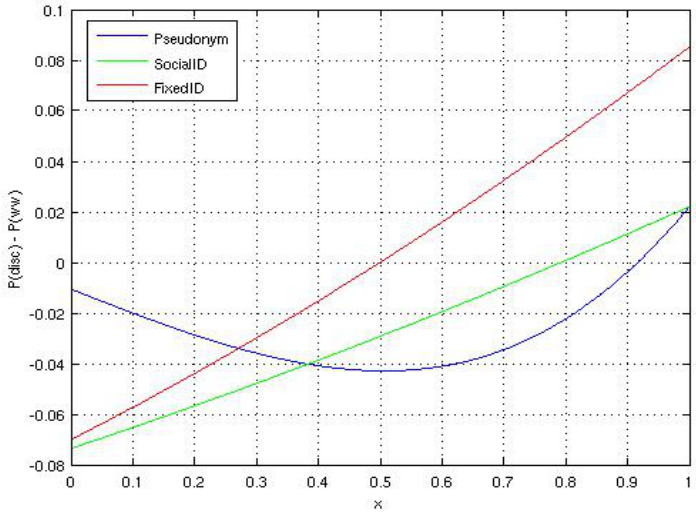


Fig. 5.4e x vs. payoff difference (p = 0.1, q = 0.9)

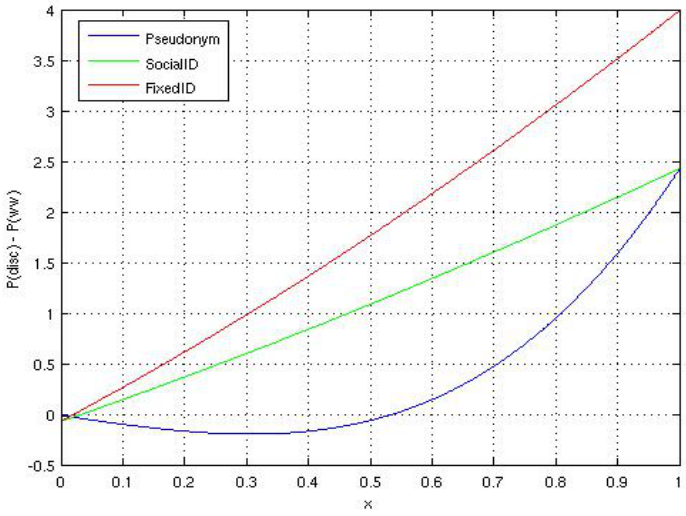


Fig. 5.4f x vs. payoff difference (b = 10, p = 0.1, q = 0.9)

than x_0. Therefore, we need more x_1 in discriminators in order to compete with whitewashers. And Figure 5.4g shows that the population of x_1 is decreasing after the jump in the first round in each generation with pseudonym. But in Figure 5.4h we can see that the population of x_1 is always growing throughout the generation. Due to this difference, even if we set $x = 0.6$ and pseudonym can defeat whitewashers, pseudonym is slower than social identity in eliminating whitewashers.

5.4.3 Cost of whitewashers

We set $x = 0.6$ so that we can compare the two identity types' cost of eliminating whitewashers. As shown in Figure 5.4i, pseudonym takes 34 generations and discriminators lose 71.8768 when whitewashers are getting eliminated. After whitewashers are gone, discriminators get 11.0561 on average for each generation. Figure 5.4j shows that social identity takes 28 generations and has a cost of 20.2680 and discriminators will get 11.0569 on average each generation once whitewashers are eliminated.

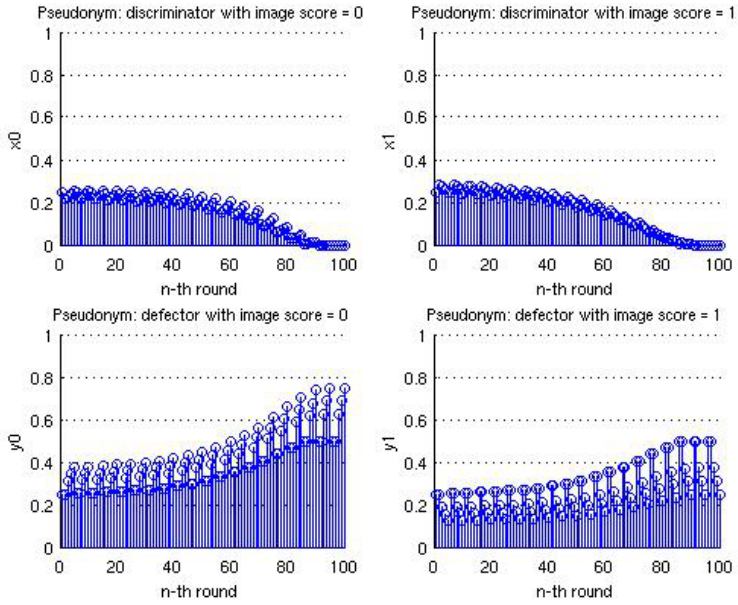


Fig. 5.4g Population Changes of Pseudonym

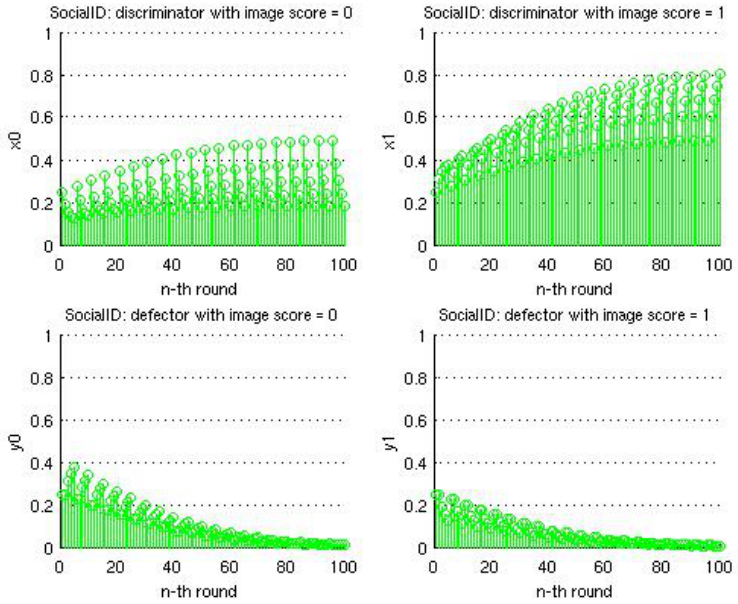


Fig. 5.4h Population Changes of Social Identity

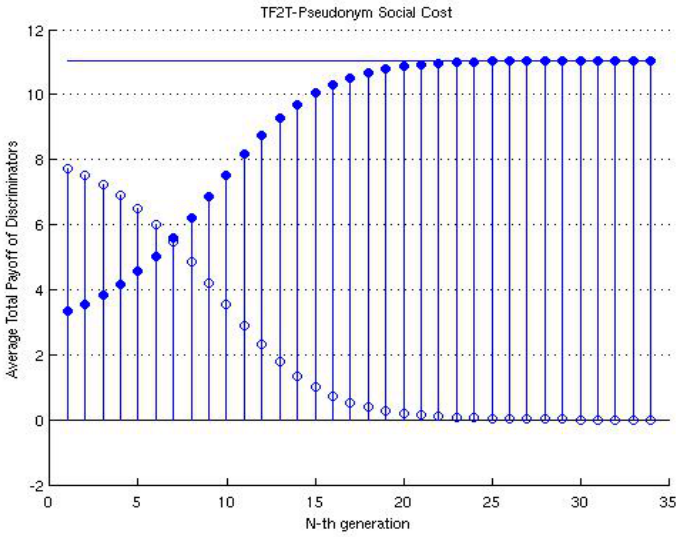


Fig. 5.4i Discriminators Payoff of Pseudonym

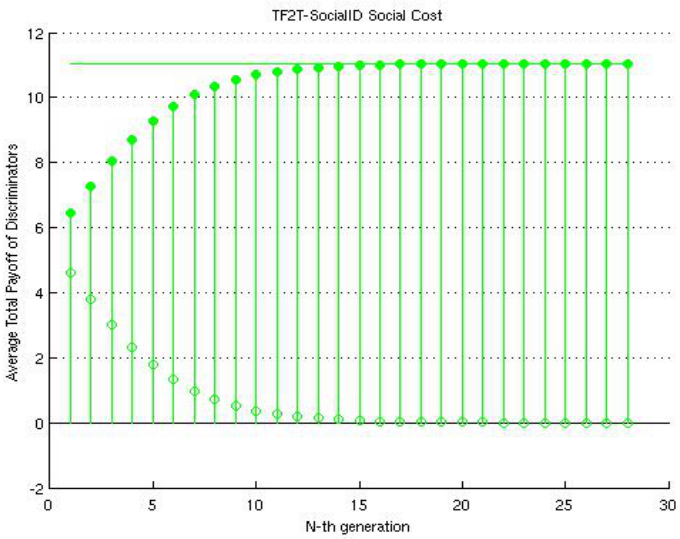


Fig. 5.4j Discriminators Payoff of Social Identity

Chapter 6

Conclusions and Future Work

6.1 Conclusions

When the community is full of whitewashers, being a whitewasher is the lucrative choice. If we do not try to create a friendly online environment, users will abandon the online community at the least. Good users will be scammed all the time. Eventually, there is no point in being a member of this online community. Therefore, we should always make an effort to protect our users from whitewashers and other type of scams.

Social identity has a consistent preference over p and q in all four strategies while pseudonym asks for different properties on p and q in image score than in other strategies. Social identity can always eliminate whitewashers faster than pseudonym and with a smaller cost. Social identity has a larger region, in terms of p, q and x, where discriminators can wipe out whitewashers. And except image score, in the other three strategies, social identity generates the same average payoff for discriminators after the whitewashers are gone. The reason that pseudonym yields more than social identity is because the p we use for pseudonym is higher and q is lower, which means more players cooperate with x_0 recipients in pseudonym than in social identity.

Notice that our definition of social cost is different from Friedman and Resnick's definition in [17]. We focus on how much discriminators lose due to whitewashers. Their definition reflects how newcomers are welcomed when first joining the community. The probability p is how often a player, whose label is unknown, will be trusted and we can use this probability as a sign for Friedman and Resnick's definition of social identity. If p is larger, more newcomers will be trusted when they first join the community. From chapter 5, we can conclude that to achieve the same payoff difference (or number of generations needed to eliminate whitewashers), social identity always has a higher, if not equal, tolerance on p than pseudonym. That is to say, veteran players are more friendly to newcomers in social identity.

In all four strategies, social identity has a preference for a larger q, which means if a social identity community is to thrive, more players need to release their information to other people. This observation might seem against many people's will to protect their privacy. However, what social identity is asking for is not publishing everything about the user to others but only the information that should be taken into consideration when the donor is deciding whether to cooperate with you. Recently, Google has made a similar announcement about it's newly launched social network product Google+ that private Google profiles will be deleted after July 31^{st}. Facebook had done the same in 2010 releasing users' names and profile pictures to non-user. All these can be seen as increasing q in our model and this would help in building a better online community.

What makes people reluctant about releasing their information on social networks is that a user alone does not see the purpose of giving his information to others. Social identity is user-centric and it's very likely that users are not aware of how his single choice to cooperate with another can affect the whole community. Each user looks at the network just from his own perspective. In our model, each discriminator simply adheres to the prescribed strategy but in real life users are not so consistent. Only when the user sees the risk of blindly cooperating with someone does he learn to avoid it. We want to build a secure neighborhood but still want our citizens to be aware of certain dangers.

As for the strategies, PYD can always generates positive payoff difference with either pseudonym or social identity and eliminate whitewashers. But it is hard to implement because it depends on the value of b, which decides how many times a user should pay for dues. In PYD, x_0 and y_0 players have to pay b dues before becoming a veteran. However, in real life users want different things from the Internet. It's hard to decide the value of b. For example, if we fix b to 10 and a user who wants to get a file that is worth only 1, there's no point for the user to pay 10 first and get just 1. In fact, there is no best strategy. Because users may follow various plans, we cannot fix a strategy for users, rather, build our online systems compatible with more strategies. No matter what strategies users are using, discriminators can still defeat whitewashers.

6.2 Future work

We analyze four strategies but other strategies do exist. Not all users would use the same strategy, future work could be done simulating multiple strategies interacting with each other. Especially after whitewashers are gone, we are interested in which strategy offers more payoff for discriminators. Also, the game we use requires only one player's decision in each round. And how much a cooperation costs as well as how much the recipient will get are fixed. Making the "prices" more flexible may unveil more about different identity types.

One of social identity's advantages is that it is more expensive than pseudonym. Future work

can take identity cost into account when calculating payoffs of each type of player. Of course, only whitewashers will spend a lot on getting new identities and thus, discriminators can get rid of whitewashers even faster with social identity than with pseudonym.

Our work could be extended to study which identity type and strategy is more suitable for a specific type of website. Because social identity puts the user at the center of everything, but not all online services is user-centric.

Appendix A

Players Payoff in PYD

A.1 Fixed Identity

$$x_0 \text{ PYD when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 1 \\ y_0 \to 0 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will get } -c \tag{A.1}$$

$$x_0 \text{ as recipient when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.2}$$

$$P(x_0) = \frac{1}{2}(-c)(x_1 + y_1) + \frac{1}{2}by_0 \tag{A.3}$$

$$x_1 \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 1 \\ y_0 \to 0 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will get } -c \tag{A.4}$$

$$x_1 \text{ as recipient when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.5}$$

$$P(x_1) = \frac{1}{2}(-c)(x_1 + y_1) + \frac{1}{2}b(x + y_0) \tag{A.6}$$

$$y_0 \text{ PYD when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will get } -c \tag{A.7}$$

$$y_0 \text{ as recipient when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.8}$$

$$P(y_0) = \frac{1}{2}(-c) + \frac{1}{2}by_0 \tag{A.9}$$

$$y_1 \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{A.10}$$

$$y_1 \text{ as recipient when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.11}$$

$$P(y_1) = \frac{1}{2}b(x + y_0) \tag{A.12}$$

A.2 Pseudonym

$$x_0 \text{ PYD when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to q(1-y) + (1-q)p \\ y_0 \to (1-q)p \\ y_1 \to q(1-y) + (1-q)p \end{cases} \Rightarrow \text{will get } -c \tag{A.13}$$

$$x_0 \text{ as recipient when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 \to (1-q)p \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.14}$$

$$P(x_0) = \frac{1}{2}(-c)\left[(1-q)p + (x_1+y_1)q(1-y)\right] + \frac{1}{2}b\left[x(1-q)p + y_0\right] \tag{A.15}$$

$$x_1 \text{ as donor when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 \to q(1-y) + (1-q)p \\ y_0 \to (1-q)p \\ y_1 \to q(1-y) + (1-q)p \end{cases} \Rightarrow \text{will get } -c \tag{A.16}$$

$$x_1 \text{ as recipient when meeting} \begin{cases} x_0 \to q(1-y) + (1-q)p \\ x_1 \to q(1-y) + (1-q)p \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.17}$$

$$P(x_1) = \frac{1}{2}(-c)\left[x(1-q)p + x_1q(1-y) + py\right] + \frac{1}{2}b\left[x(q(1-y) + (1-q)p) + y_0\right] \tag{A.18}$$

$$y_0 \text{ PYD when meeting} \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will get } -c \tag{A.19}$$

$$y_0 \text{ as recipient when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 \to (1-q)p \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.20}$$

$$P(y_0) = \frac{1}{2}(-c) + \frac{1}{2}b\left[x(1-q)p + y_0\right] \tag{A.21}$$

$$y_1 \text{ as donor when meeting} \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{A.22}$$

$$y_1 \text{ as recipient when meeting} \begin{cases} x_0 \to q(1-y) + (1-q)p \\ x_1 \to q(1-y) + (1-q)p \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.23}$$

$$P(y_1) = \frac{1}{2}b(xp + y_0) \tag{A.24}$$

A.3 Social Identity

Payoff of each type of player's:

$$x_0 \text{ PYD when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 \to q + (1-q)p \\ y_0 \to (1-q)p \\ y_1 \to q + (1-q)p \end{cases} \Rightarrow \text{will get } -c \tag{A.25}$$

$$x_0 \text{ as recipient when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 \to (1-q)p \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.26}$$

$$P(x_0) = \frac{1}{2}(-c)[(1-q)p + (x_1 + y_1)q] + \frac{1}{2}b[x(1-q)p + y_0] \tag{A.27}$$

$$x_1 \text{ as donor when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 \to q + (1-q)p \\ y_0 \to (1-q)p \\ y_1 \to p \end{cases} \Rightarrow \text{will get } -c \tag{A.28}$$

$$x_1 \text{ as recipient when meeting} \begin{cases} x_0 \to q + (1-q)p \\ x_1 \to q + (1-q)p \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.29}$$

$$P(x_1) = \frac{1}{2}(-c)[(x + y_0)(1-q)p + x_1 q + p y_1] + \frac{1}{2}b[x(q + (1-q)p) + y_0] \tag{A.30}$$

$$y_0 \text{ PYD when meeting} \begin{cases} x_0 \to 1 \\ x_1 \to 1 \\ y_0 \to 1 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will get } -c \tag{A.31}$$

$$y_0 \text{ as recipient when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 \to (1-q)p \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.32}$$

$$P(y_0) = \frac{1}{2}(-c) + \frac{1}{2}b[x(1-q)p + y_0] \tag{A.33}$$

$$y_1 \text{ as donor when meeting} \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{A.34}$$

$$y_1 \text{as recipient when meeting} \begin{cases} x_0 \to q + (1-q)p \\ x_1 \to q + (1-q)p \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{A.35}$$

$$P(y_1) = \frac{1}{2} b(xp + y_0) \tag{A.36}$$

Appendix B

Players Payoff in Pavlov

B.1 Fixed Identity

Payoff of each type of player's can be calculated as follow:

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 0 \\ y_0 \to 1 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{x_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(x_0) = \frac{1}{2}(-c)(x_0 + y_0) + \frac{1}{2}bx_0$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 1 \\ y_0 \to 0 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{x_1}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 1 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(x_1) = \frac{1}{2}(-c)(x_1 + y_1) + \frac{1}{2}bx_1$$

$$\frac{y_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{y_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to 1 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(y_0) = \frac{1}{2}bx_0$$

$$\frac{y_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \text{ will get } -c$$

$$\frac{y_1}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 1 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(y_1) = \frac{1}{2}bx_1$$

B.2 Pseudonym

Payoff of each type of player's can be calculated as follow:

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q(1-y)+(1-q)p \\ x_1 \to (1-q)p \\ y_0 \to p \\ y_1 \to (1-q)p \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{x_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to q(1-y)+(1-q)p \\ x_1 \to (1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(x_0) = \frac{1}{2}(b-c)\left[(q(1-y)+(1-q)p)x_0(k) + x_1(1-q)p\right] + \frac{1}{2}(-c)\left[y_0(k)p + p(1-q)y_1(k)\right]$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to q(1-y)+(1-q)p \\ y_0 \to p \\ y_1 \to q(1-y)+(1-q)p \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{x_1}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to q(1-y)+(1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$\begin{aligned} P(x_1) =& \frac{1}{2}(b-c)\left[x_0(k)(1-q)p + x_1(k)(q(1-y)+(1-q)p)\right] \\ &+ \frac{1}{2}(-c)\left[py_0(k) + (q(1-y)+(1-q)p)y_1(k)\right] \end{aligned}$$

$$\frac{y_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{y_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to p \\ x_1 \to p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(y_0) = \frac{1}{2}bpx$$

$$\frac{y_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{y_1}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to q(1-y)+(1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(y_1) = \frac{1}{2}b[(1-q)px + q(1-y)x_1(k)]$$

B.3 Social Identity

Payoff of each type of player's can be calculated as follow:

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to q+(1-q)p \\ x_1 \to (1-q)p \\ y_0 \to p \\ y_1 \to (1-q)p \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{x_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to q+(1-q)p \\ x_1 \to (1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(x_0) = \frac{1}{2}(b-c)[qx_0(k) + x(1-q)p] + \frac{1}{2}(-c)[y_0(k)p + p(1-q)y_1(k)]$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to q + (1-q)p \\ y_0 \to p \\ y_1 \to q + (1-q)p \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{x_1}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to q + (1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(x_1) = \frac{1}{2}(b-c)[x_0(k)(1-q)p + x_1(k)(q+(1-q)p)] + \frac{1}{2}(-c)[py_0(k) + (q+(1-q)p)y_1(k)]$$

$$\frac{y_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{y_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to p \\ x_1 \to p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(y_0) = \frac{1}{2}bpx$$

$$\frac{y_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c$$

$$\frac{y_1}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 \to q + (1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b$$

$$P(y_1) = \frac{1}{2}b[(1-q)px + qx_1(k)]$$

Appendix C

Players Payoff in TF2T

C.1 Fixed Identity

And payoff of each type of players:

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 + z \to 1 \\ y_0 \to 0 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will get } -c \tag{C.1}$$

$$\frac{x_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to 0 \\ x_1 + z \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.2}$$

$$P(x_0) = \frac{1}{2}(-c)(x_1 + z + y_1) \tag{C.3}$$

$$\frac{x_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 + z \to 1 \\ y_0 \to 0 \\ y_1 \to 1 \end{cases} \Rightarrow \text{will get } -c \tag{C.4}$$

$$\frac{x_1}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to 1 \\ x_1 + z \to 1 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.5}$$

$$P(x_1) = \frac{1}{2}(-c)(x_1 + z + y_1) + \frac{1}{2}b(x_0 + x_1 + z) \tag{C.6}$$

$$\frac{y_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{C.7}$$

$$\frac{y_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.8}$$

$$P(y_0) = 0 \tag{C.9}$$

$$\frac{y_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1 \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{C.10}$$

$$\frac{y_1}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to 1 \\ x_1 + z \to 1 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.11}$$

$$P(y_1) = \frac{1}{2}b(x_0 + x_1 + z) \tag{C.12}$$

C.2 Pseudonym

Payoff of each type of player's:

$$\frac{x_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 + z \to q(1-y) + (1-q)p \\ y_0 \to p \\ y_1 \to q(1-y) + (1-q)p \end{cases} \Rightarrow \text{will get } -c \tag{C.13}$$

$$\frac{x_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1 + z \to (1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.14}$$

$$P(x_0) = \frac{1}{2}(-c)[(x+y_1)(1-q)p + (x_1+z+y_1)q(1-y) + py_0] + \frac{1}{2}b[x(1-q)p] \tag{C.15}$$

$$\frac{x_1+z}{2} \text{ as donor when meeting } \begin{cases} x_0 \to (1-q)p + q(1-y) \\ x_1 + z \to q(1-y) + (1-q)p \\ y_0 \to p \\ y_1 \to q(1-y) + (1-q)p \end{cases} \Rightarrow \text{will get } -c \tag{C.16}$$

$$\frac{x_1+z}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to q(1-y) + (1-q)p \\ x_1 + z \to q(1-y) + (1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.17}$$

$$P(x_1) = \frac{1}{2}(-c)[(x+y_1)(1-q)p + (x_1+z+y_1)q(1-y) + py_0] + \frac{1}{2}bx[q(1-y) + (1-q)p] \tag{C.18}$$

$$P(z) = P(x_1) \tag{C.19}$$

$$\frac{y_0}{2} \text{ as donor when meeting} \begin{cases} x_0 \to 0 \\ x_1 + z \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{C.20}$$

$$\frac{y_0}{2} \text{ as recipient when meeting} \begin{cases} x_0 \to p \\ x_1 + z \to p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.21}$$

$$P(y_0) = \frac{1}{2}bxp \tag{C.22}$$

$$\frac{y_1}{2} \text{ as donor when meeting} \begin{cases} x_0 \to 0 \\ x_1 + z \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{C.23}$$

$$\frac{y_1}{2} \text{ as recipient when meeting} \begin{cases} x_0 \to q(1-y) + (1-q)p \\ x_1 + z \to q(1-y) + (1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.24}$$

$$P(y_1) = \frac{1}{2}bx[q(1-y) + (1-q)p] \tag{C.25}$$

C.3 Social Identity

Payoff of each type of player's:

$$\frac{x_0}{2} \text{ as donor when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 + z \to q + (1-q)p \\ y_0 \to p \\ y_1 \to q + (1-q)p \end{cases} \Rightarrow \text{will get } -c \tag{C.26}$$

$$\frac{x_0}{2} \text{ as recipient when meeting} \begin{cases} x_0 \to (1-q)p \\ x_1 + z \to (1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.27}$$

$$P(x_0) = \frac{1}{2}(-c)[(x + y_1)(1-q)p + (x_1 + z + y_1)q + py_0] + \frac{1}{2}b[x(1-q)p] \tag{C.28}$$

$$\frac{x_1+z}{2} \text{ as donor when meeting } \begin{cases} x_0 \to (1-q)p \\ x_1+z \to q+(1-q)p \\ y_0 \to p \\ y_1 \to q+(1-q)p \end{cases} \Rightarrow \text{will get } -c \tag{C.29}$$

$$\frac{x_1+z}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to q+(1-q)p \\ x_1+z \to q+(1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.30}$$

$$P(x_1) = \frac{1}{2}(-c)[(x+y_1)(1-q)p + (x_1+z+y_1)q + py_0] + \frac{1}{2}bx[q+(1-q)p] \tag{C.31}$$

$$P(z) = P(x_1) \tag{C.32}$$

$$\frac{y_0}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1+z \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{C.33}$$

$$\frac{y_0}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to p \\ x_1+z \to p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.34}$$

$$P(y_0) = \frac{1}{2}bxp \tag{C.35}$$

$$\frac{y_1}{2} \text{ as donor when meeting } \begin{cases} x_0 \to 0 \\ x_1+z \to 0 \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } -c \tag{C.36}$$

$$\frac{y_1}{2} \text{ as recipient when meeting } \begin{cases} x_0 \to q+(1-q)p \\ x_1+z \to q+(1-q)p \\ y_0 \to 0 \\ y_1 \to 0 \end{cases} \Rightarrow \text{will get } b \tag{C.37}$$

$$P(y_1) = \frac{1}{2}bx(q+(1-q)p) \tag{C.38}$$

References

[1] G. Akerlof and R. Kranton, "Economics and Identity," *Quarterly journal of Economics*, vol. 115, no. 3, pp. 715–753, 2000.

[2] B. Eaton, M. Eswaran, and R. Oxoby, "'Us' and 'Them': The Origin of Identity, and Its Economic Implications," *Working Papers*, 2009.

[3] D. Chaum, "Security Without Identification: Transaction Systems to Make Big Brother Obsolete," *Communications of the ACM*, vol. 28, no. 10, pp. 1030–1044, 1985.

[4] D. Recordon and D. Reed, "Openid 2.0: a platform for user-centric identity management," in *Proceedings of the second ACM workshop on Digital identity management*, pp. 11–16, ACM, 2006.

[5] H. Oh and S. Jin, "The security limitations of sso in openid," in *Advanced Communication Technology, 2008. ICACT 2008. 10th International Conference on*, vol. 3, pp. 1608–1611, IEEE, 2008.

[6] T. Neubauer and B. Riedl, "Improving patients privacy with pseudonymization.," *Studies in health technology and informatics*, vol. 136, p. 691, 2008.

[7] F. Stutzman, "An evaluation of identity-sharing behavior in social network communities," *International Digital and Media Arts Journal*, vol. 3, no. 1, pp. 10–18, 2006.

[8] L. Sweeney *et al.*, "k-anonymity: A model for protecting privacy," *International Journal of Uncertainty Fuzziness and Knowledge Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[9] A. Pashalidis and B. Meyer, "Linking anonymous transactions: The consistent view attack," in *Privacy Enhancing Technologies*, pp. 384–392, Springer, 2006.

[10] G. Chen and F. Rahman, "Analyzing privacy designs of mobile social networking applications," in *Embedded and Ubiquitous Computing, 2008. EUC'08. IEEE/IFIP International Conference on*, vol. 2, pp. 83–88, IEEE, 2008.

[11] A. Pashalidis and C. Mitchell, "Limits to anonymity when using credentials," in *Security Protocols*, pp. 4–12, Springer, 2006.

[12] H. Krasnova, O. Günther, S. Spiekermann, and K. Koroleva, "Privacy concerns and identity in online social networks," *Identity in the Information Society*, vol. 2, no. 1, pp. 39–63, 2009.

[13] A. Marshall and B. Tompsett, "Identity theft in an online world," *Computer Law & Security Report*, vol. 21, no. 2, pp. 128–137, 2005.

[14] J. Bonneau, J. Anderson, R. Anderson, and F. Stajano, "Eight friends are enough: social graph approximation via public listings," in *Proceedings of the Second ACM EuroSys Workshop on Social Network Systems*, pp. 13–18, ACM, 2009.

[15] J. He, W. Chu, and Z. Liu, "Inferring privacy information from social networks," *Intelligence and Security Informatics*, pp. 154–165, 2006.

[16] L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda, "All your contacts are belong to us: automated identity theft attacks on social networks," in *Proceedings of the 18th international conference on World wide web*, pp. 551–560, ACM, 2009.

[17] E. Friedman and P. Resnick, "The Social Cost of Cheap Pseudonyms," *Ann Arbor*, vol. 1001, pp. 48109–1092, 1999.

[18] E. Adar and B. Huberman, "Free Riding on Gnutella," *First Monday*, vol. 5, no. 10, pp. 2–13, 2000.

[19] M. Feldman, C. Papadimitriou, J. Chuang, and I. Stoica, "Free-Riding and Whitewashing in Peer-to-Peer Systems," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 5, pp. 1010–1019, 2006.

[20] B. Levine, C. Shields, and N. Margolin, "A survey of solutions to the sybil attack," *University of Massachusetts Amherst, Amherst, MA*, 2006.

[21] J. Douceur, "The sybil attack," *Peer-to-peer Systems*, pp. 251–260, 2002.

[22] H. Yu, M. Kaminsky, P. Gibbons, and A. Flaxman, "Sybilguard: defending against sybil attacks via social networks," *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 4, pp. 267–278, 2006.

[23] A. Lysyanskaya, R. Rivest, A. Sahai, and S. Wolf, "Pseudonym systems," in *Selected Areas in Cryptography*, pp. 184–199, Springer, 2000.

[24] P. Penna, F. Schoppmann, R. Silvestri, and P. Widmayer, "Pseudonyms in cost-sharing games," *Internet and Network Economics*, pp. 256–267, 2009.

[25] B. Ford and J. Strauss, "An offline foundation for online accountable pseudonyms," in *Proceedings of the 1st workshop on Social network systems*, pp. 31–36, ACM, 2008.

[26] M. Yokoo, Y. Sakurai, and S. Matsubara, "The effect of false-name bids in combinatorial auctions: new fraud in internet auctions* 1," *Games and Economic Behavior*, vol. 46, no. 1, pp. 174–188, 2004.

[27] J. Brainard, A. Juels, R. Rivest, M. Szydlo, and M. Yung, "Fourth-Factor Authentication: Somebody You Know," in *Proceedings of the 13th ACM conference on Computer and communications security*, pp. 168–178, ACM, 2006.

[28] B. Soleymani and M. Maheswaran, "Social authentication protocol for mobile phones," in *Computational Science and Engineering, 2009. CSE'09. International Conference on*, vol. 4, pp. 436–441, IEEE, 2009.

[29] B. Viswanath, A. Post, K. Gummadi, and A. Mislove, "An analysis of social network-based sybil defenses," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 4, pp. 363–374, 2010.

[30] L. Jin, H. Takabi, and J. Joshi, "Towards active detection of identity clone attacks on online social networks," in *Proceedings of the first ACM conference on Data and application security and privacy*, pp. 27–38, ACM, 2011.

[31] A. Mislove, B. Viswanath, K. Gummadi, and P. Druschel, "You are who you know: Inferring user profiles in online social networks," in *Proceedings of the third ACM international conference on Web search and data mining*, pp. 251–260, ACM, 2010.

[32] D. Watts, P. Dodds, and M. Newman, "Identity and search in social networks," *Science*, vol. 296, no. 5571, p. 1302, 2002.

[33] A. Mislove, M. Marcon, K. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, pp. 29–42, ACM, 2007.

[34] M. Subramani and B. Rajagopalan, "Knowledge-sharing and influence in online social networks via viral marketing," *Communications of the ACM*, vol. 46, no. 12, pp. 300–307, 2003.

[35] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Predicting positive and negative links in online social networks," in *Proceedings of the 19th international conference on World wide web*, pp. 641–650, ACM, 2010.

[36] L. Kimball and H. Rheingold, "How online social networks benefit organizations," *Rheingold Associates*, 2000.

[37] J. Mankoff, D. Matthews, S. Fussell, and M. Johnson, "Leveraging social networks to motivate individuals to reduce their ecological footprints," 2007.

[38] R. Cachia, R. Compañó, and O. Da Costa, "Grasping the potential of online social networks for foresight," *Technological Forecasting and Social Change*, vol. 74, no. 8, pp. 1179–1203, 2007.

[39] M. Maheswaran, B. Ali, H. Ozguven, and J. Lord, "Online identities and social networking," *Handbook of Social Network Technologies and Applications*, pp. 241–267, 2010.

[40] H. Choi, S. Kruk, S. Grzonkowski, K. Stankiewicz, B. Davis, and J. Breslin, "Trust Models for Community-Aware Identity Management," in *Identity, Reference, and the Web Workshop (IRW 2006)*, Citeseer, 2006.

[41] B. Jennings and A. Finkelstein, "Digital Identity and Reputation in The Context of A Bounded Social Ecosystem," in *Business Process Management Workshops*, pp. 687–697, Springer, 2009.

[42] N. Agarwal, H. Liu, S. Murthy, A. Sen, and X. Wang, "A social identity approach to identify familiar strangers in a social network," in *Proceedings of the 3rd International AAAI Conference of Weblogs and Social Media*, 2009.

[43] B. Bruin, "Game Theory in Philosophy," *Topoi*, vol. 24, no. 2, pp. 197–208, 2005.

[44] T. Webster, *Introduction to Game Theory in Business and Economics*. M.E. Sharpe, 2009.

[45] P. Goodwin, "Forecasting Games: Can Game Theory Win?," *International Journal of Forecasting*, vol. 18, no. 3, pp. 369–374, 2002.

[46] L. Amsel, A. Pilpel, and R. Marshal, "Towards A Mathematical Psychiatry Game Theory Modeling Of OCD," *Association for the Scientific Study of Consciousnes 2007, ASSC11*, 2006.

[47] R. Fisher, *Theory of Natural Selection*. Oxford University Press, London, 1930.

[48] L. Samuelson, *Evolutionary Games and Equilibrium Selection*, vol. 1. The MIT press, 1998.

[49] J. Hofbauer and K. Sigmund, *Evolutionary games and population dynamics*. Cambridge Univ Pr, 1998.

[50] J. Watson, *Strategy: An Introduction to Game Theory*. Pergamon Press., 2001.

[51] M. Feldman and J. Chuang, "The Evolution of Cooperation Under Cheap Pseudonyms," in *E-Commerce Technology, 2005. CEC 2005. Seventh IEEE International Conference on*, pp. 284–291, IEEE, 2005.

[52] R. Axelrod and W. Hamilton, "The Evolution of Cooperation," *Science*, vol. 211, no. 4489, p. 1390, 1981.

[53] M. Nowak and K. Sigmund, "Evolution of indirect reciprocity by image scoring," *Nature*, vol. 393, no. 6685, pp. 573–577, 1998.

[54] M. Nowak and K. Sigmund, "Tit for tat in heterogeneous populations," *Nature*, vol. 355, no. 6357, pp. 250–253, 1992.