An Extension of the Mafia Game Analysis

*Lluc Puig-Codina**

January 30, 2018

Abstract

This note extends the existing game theoretical and probabilistic analysis of the mafia game to the (one) *kamikaze* variation from it's most simple version, n citizens and one mobster. Both the probability of the mafia winning the game and the optimal stoppage rule for the *kamikaze* are calculated.

Keywords: Mafia game; Perfect play; optimal stopping.

Introduction

Everyone has played the Mafia game (also called Werewolf) at some point, a classic game of an informed minority *versus* an uninformed majority. In it's simplest form, a group of $n + 1 > 3$ people is secretly divided into two groups, a group of size $c \equiv n - m$, citizens, and a group of size m, the mafia. The remaining person is the organizer. The members of the mafia know which members belong in their group, while citizens don't. The game operates in a sequential fashion, turn after turn, and within a turn there is first, day, and later, night. During the day, all n players vote on the member they will lynch, and during the night the mob votes on who they will assasinate. The game is won when only the members of one group remain alive. From a game theoretic perspective, if players could perfectly conceal their

*llucpuigcodina@gmail.com

behaviour (say some psychological factors are ommited here) it's trivial to sense that lyinchings and assasinations should be random. Braverman, Etesami and Mossel (2008) prove this intuition by showing it is indeed a Nash equilibria.

The previously mentioned authors also show that for a game to be asymptotically *fair*, in the simple setting, there should be a mafia size of order $\sqrt{n}$. Unfortunately their theorem can't guarantee such solution albeit it is proven to be correct, see Yao (2008). In the case of the detective variation, for the game to be asymptotically *fair*, there should be a mafia size of order n. Piotr Migdal (2013) finds the close form solution and analyses the dynamic properties of Braverman, Etesami and Mossel's model.

In the particular case of the simple game with only one mafia member, the probability of the mafia winning is given by:

$$\omega(n,1) = \frac{n-1}{n} * \frac{n-3}{n-2} * ... * \frac{1+(n \mod 2)}{2+(n \mod 2)} = \prod_{i=0}^{[\frac{n}{2}]-1} \left(\frac{n-(2i+1)}{n-2i} \right) = \frac{(n-1)!!}{n!!} \tag{1}$$

Note that in the case where only one citizen is left, the lyinching is decided by a fair coin toss.

Kamikaze

The kamikaze is a recent innovation that introduces interesting endgame dynamics. The kamikaze is a role a citizen can take, and it can activate itself to kill both himself and a chosen player[1], without regard for whom the other players choose to lynch. See figure 1 for an example. In our study we will assume that the kamikaze

[1] https://wiki.mafiascum.net/index.php?title=Suicide_Bomber

can activate itself each turn, right after the citizens lynch (and the identity of the lynched is revealed) and before the mafia assasinates. This is done both to avoid the complications of a kamikaze conditioning the lynching and viceversa, and to give a, hopefully interesting, stategic advantage to the kamikaze. The justification for this might be that of the behaviour of a radical who disagrees with the democratic outcome of the players and decides to take justice into his own hands.

To see the importance of the kamikaze in the endgame, say that that we reach a point in which there are 2 citizens left, one of them a kamikaze, and one mobster. Assume that the citizen is picked to be lynched. Then, the kamikaze could activate it's ability and kill the mobster, forcing a draw decided by a fair coin toss, in what would otherwise be a loss. If the mobster is picked to be lynched, citizens win, and if the kamikaze is picked the mafia wins. Introducing a single kamikaze changes $\omega(3,1)$ from $2/3$ to $(1/3) + (1/3) * (1/2) = 1/2$.

Figure 1: Kamikaze[2]

Our first question is: *under what conditions it is optimal for the kamikaze to activate?* Let k denote the kamikaze. From any state $(n, k = 1, m)$:

[2] https://wiki.mafiascum.net/images/f/fb/M-suicidebomber.png

$$P_k(killing\ a\ mobster) = \frac{m}{n-1}$$

$$P_k(killing\ a\ citizen) = \frac{n-m-1}{n-1}$$

Thus, for $m = 1$, it is optimal to activate the kamikaze when (assume the game is set up such that comunicating one's identity is cheap talk, therefore random play is still the optimal strategy for all players during the lynching and assasination, which is the case due to the signal being costless, the kamikaze's incentives not to reveal his true identity and others not to believe him when he does so):

$$\textit{Optimal Stoppage Condition:} \qquad \underbrace{K}_{P(\text{mafia win}|\text{kamikaze active in the current state})} \quad < \quad \Psi(n,1,1)$$

where $\Psi(n,k,m)$ is the probability that the mafia wins if the kamikaze does not become active in the current round, assuming kamikaze optimal play from there on.

$$K(n,1,1) = \frac{1}{n} * 0 + \frac{1}{n} * \omega(n-2,1) + \frac{n-2}{n} * \left(\frac{n-3}{n-2} * \omega(n-4,1) + \frac{1}{n-2} * 0 \right)$$

where first term of *K(n,1,1)* corresponds to the chance of lynching the mobster in the current state, and the second term corresponds to the chance of lynching the kamikaze, in which case the game colapses to the simple game with 2 players less. The third term of *K(n,1,1)* first corresponds to the probability of lynching an ordinary citizen in the current state, followed by the probability that the kamikaze kills a citizen, in which case the game colapses to the simple game with 4 players less (the mobster will them assasinate one citizen), and the probability of the kamikaze killing the mobster.

This optimal stoppage condition allows us to define the kamikaze optimal play probability at each period:

$$\phi(n,1,1) = \min\{K(n,1,1), \Psi(n,1,1)\} \qquad (2)$$

Note that $\phi(n,0,1) = \Psi(n,0,1) = \omega(n,1)$, and $\phi(n,1,0) = \phi(n,0,0) = \Psi(n,1,0) = \psi(n,0,0) = \omega(n,0) = 0$.

Now, we can go on to answer a second question: *assuming a kamikaze optimal play, what is the probability that the mafia wins?* We start by defining $\Psi(n,1,1)$ in terms of the kamikaze optimal play condition and the standard mafia simple game:

$$\begin{aligned} \Psi(n,1,1) = & \underbrace{\frac{1}{n}}_{3.1} *\omega(n-2,1) \\ + & \underbrace{\frac{1}{n}}_{3.2} *0 \\ + & \underbrace{\left(\frac{n-2}{n} * \frac{1}{n-1}\right)}_{3.3} *\omega(n-2,1) \\ + & \underbrace{\left(\frac{n-2}{n} * \frac{n-3}{n-2}\right)}_{3.4} *\phi(n-2,1,1) \end{aligned} \qquad (3)$$

where 3.1 corresponds to the probability of the kamikaze being lynched, followed by the probability that the mafia win in such case. 3.2 corresponds to the probability of the mobster being lynched, followed by the proability of the mafia winning in such case. 3.3 is the probability of an ordinary citizen being lynched and the kamikaze being assasinated, followed by the probability that the mafia win in such case. 3.4

corresponds to the probability of an ordinary citizen being lynched and another ordinady citizen being assasinated, followed by the probability of the mafia winning in such case, taking into account the kamikaze's optimal play from there on.

Finally, we can solve the recursive system given by 1, 2 and 3 backwards, by imposing an end state condition:

$$\textit{End State Condition:} \begin{cases} \phi(2,1,1) = \Psi(2,1,1) = \frac{1}{2} \\ \phi(3,1,1) = \frac{1}{2} \\ \Psi(3,1,1) = \frac{2}{3} \\ \phi(4,1,1) = \Psi(4,1,1) = \frac{6}{16} \end{cases} \tag{4}$$

At last, we can now answer the second question, finding the values of $\phi(n,1,1)$ in Figure 2, by going backwards from the end state condition.

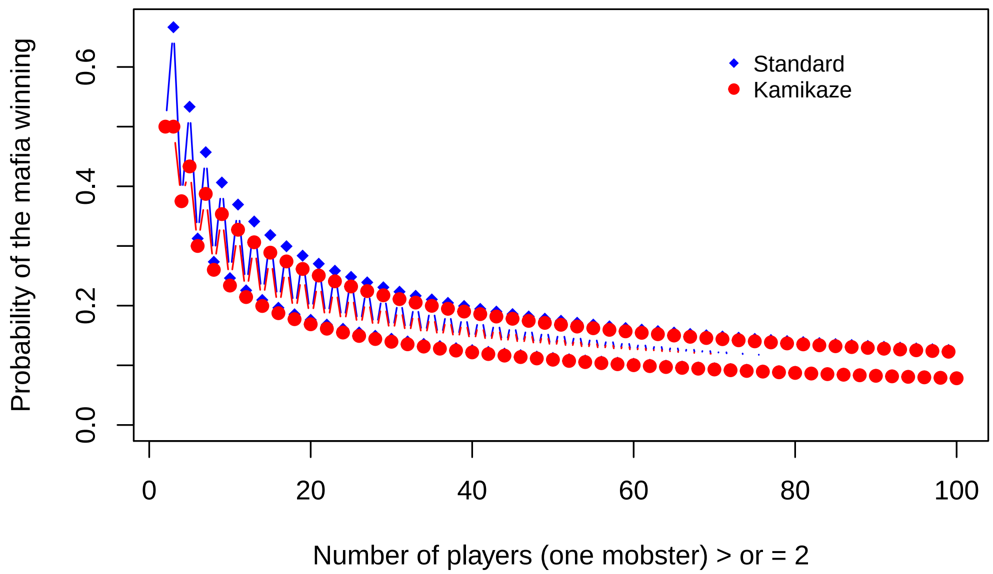

Figure 2: Kamikaze versus Standard

In figure 3, we can observe the values of $K(n1,1,)$ and $\Psi(n,1,1)$ which indicate the periods at which it is optimal for the kamikaze to activate. If $K(n1,1,) < \Psi(n,1,1)$ it is optimal to activate, such as in the case of $n = 3$, the last round when we start with an odd number of players. The kamikaze is also indiferent when $K(n1,1,) = \Psi(n,1,1)$, which is the case for $n = 2$ and $n = 4$, the last and penultimate round when starting with an even number of players. In all other cases, $K(n1,1,) > \Psi(n,1,1)$, it is optimal for the kamikaze to wait.

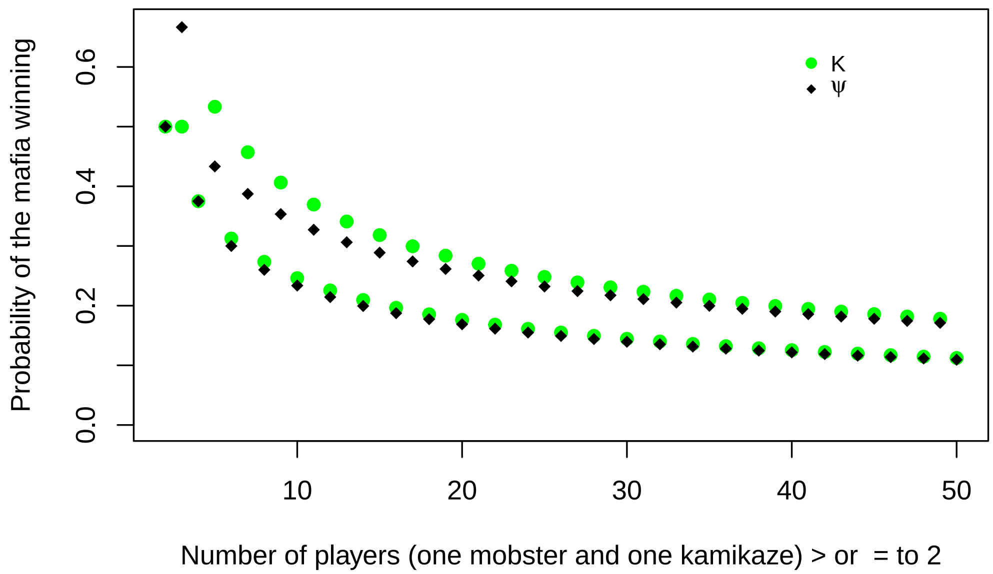

Figure 3: Optimal Stoppage

Disclaimer

Do not use against the actual mafia.

References

Braverman, M., Etesami, O., & Mossel, E. (2008). Mafia: A theoretical study of players and coalitions in a partial information environment. *The Annals of Applied Probability, 18(3), 825-846.*

Migdał, P. (2010). A mathematical model of the Mafia game. *arXiv preprint arXiv:1009.1031.*

Yao, E. (2008). A theoretical study of mafia games. *arXiv preprint arXiv:0804.0071.*

Code (in **R***)*

```
# Double Factorial Function

doublefactorial <- function(a){


        if ( a %% 2 == 0 & a %% 1 == 0 & a >= 1){


                k <- a/2

                result <- (2^k)*factorial(k)

                return(result)


        }


        if ( a %% 2 != 0 & a %% 1 == 0 & a >= 3){


                k <- (a+1)/2

                result <- factorial(2*k)/((2^k)*factorial(k))

                return(result)


        }


        if ( a == 1 ){


                return(1)


        }
```

```r
        if ( a == 0 ){

                return(1)

        }

        if ( a %% 1 != 0 | a < 0){

                print("Error: Please insert a positive integer as input.")

        }

}

#Prob of mob winning in a simple Mafia Game
w <- function(n){

        prob.win <- doublefactorial(n-1)/doublefactorial(n)
        return(prob.win)

}


#Prob of winning in a simple Mafia Game with a kamikaze
```

```
o <- function(z){

        psi <- c()
        phi <- c()
        k <- c()

        if ( z%%2 != 0){

                phi[3] <- 1/2
                psi[3] <- 2/3

                for (i in 3:( (z+1)/2 ) ){
                        n <- 2*i - 1
                        k[n] <- (1/n)*w(n-2) + ((n-2)/n)*((n-3)/(n-2))*w(n-4)
                        psi[n] <- (1/n)*w(n-2) + ((n-2)/n)*(1/(n-1))*w(n-2) + ((n-2)/n)*(
                        phi[n] <- min( k[n], psi[n])

                }

                return( phi[z] )

        }else{
                phi[2] <- 1/2
                psi[2] <- 1/2
                phi[4] <- 6/16
                psi[4] <- 6/16
```

```
                for (i in 3:( (z)/2 ) ){

                        n <- 2*i

                        k[n] <- (1/n)*w(n-2) + ((n-2)/n)*((n-3)/(n-2))*w(n-4)

                        psi[n] <- (1/n)*w(n-2) + ((n-2)/n)*(1/(n-1))*w(n-2) + ((n-2)/n)*(

                        phi[n] <- min( k[n], psi[n])


                }


                return( phi[z] )


        }


}


# Figure 1


N <- 2:100

simple <- c()

for ( i in N){

        simple[i] <- w(i)

}

N <- 5:100

kamikaze <- c()

for (i in N){
```

```
        kamikaze[i] <- o(i)
}
kamikaze[2] <-0.5
kamikaze[3] <-0.5
kamikaze[4] <- 6/16

N <- 2:100
plot(N, simple[2:100], type = "b", pch = 18, col = "blue",  xlab = "Number of players (one
points(N,kamikaze[2:100], type = "b", pch = 19, col = "red")
legend(70, 0.65, legend=c("Standard", "Kamikaze"),
       col=c("blue", "red"), pch=c(18,19), cex=0.8, box.lty=0)

# Figure 2

k <- function(z){

        psi <- c()
        phi <- c()
        k <- c()

        if ( z%%2 != 0){

                phi[3] <- 1/2
                psi[3] <- 2/3

                for (i in 3:( (z+1)/2 ) ){
```

```
                n <- 2*i - 1
                k[n] <- (1/n)*w(n-2) + ((n-2)/n)*((n-3)/(n-2))*w(n-4)
                psi[n] <- (1/n)*w(n-2) + ((n-2)/n)*(1/(n-1))*w(n-2) + ((n-2)/n)*(
                phi[n] <- min( k[n], psi[n])

        }

        return( k[z] )

}else{
        phi[2] <- 1/2
        psi[2] <- 1/2
        phi[4] <- 6/16
        psi[4] <- 6/16



        for (i in 3:( (z)/2 ) ){
                n <- 2*i
                k[n] <- (1/n)*w(n-2) + ((n-2)/n)*((n-3)/(n-2))*w(n-4)
                psi[n] <- (1/n)*w(n-2) + ((n-2)/n)*(1/(n-1))*w(n-2) + ((n-2)/n)*(
                phi[n] <- min( k[n], psi[n])

        }

        return( k[z] )
```

```
        }

}



psi <- function(z){


        psi <- c()

        phi <- c()

        k <- c()


        if ( z%%2 != 0){


                phi[3] <- 1/2

                psi[3] <- 2/3


                for (i in 3:( (z+1)/2 ) ){

                        n <- 2*i - 1

                        k[n] <- (1/n)*w(n-2) + ((n-2)/n)*((n-3)/(n-2))*w(n-4)

                        psi[n] <- (1/n)*w(n-2) + ((n-2)/n)*(1/(n-1))*w(n-2) + ((n-2)/n)*(

                        phi[n] <- min( k[n], psi[n])


                }


                return( psi[z] )
```

```
        }else{
                phi[2] <- 1/2
                psi[2] <- 1/2
                phi[4] <- 6/16
                psi[4] <- 6/16



                for (i in 3:( (z)/2 ) ){
                        n <- 2*i
                        k[n] <- (1/n)*w(n-2) + ((n-2)/n)*((n-3)/(n-2))*w(n-4)
                        psi[n] <- (1/n)*w(n-2) + ((n-2)/n)*(1/(n-1))*w(n-2) + ((n-2)/n)*(
                        phi[n] <- min( k[n], psi[n])

                }

                return( psi[z] )

        }

}

K <- c()
N <- 5:50
for (i in N){
        K[i] <- k(i)
}
```

```
K[2] <- 1/2
K[3] <- 1/2
K[4] <- 6/16


PSI <- c()
for (i in N){
        PSI[i] <- psi(i)
}
PSI[2] <- 1/2
PSI[3] <- 2/3
PSI[4] <- 6/16
N <- 2:50
plot(N, K[2:50], type = "p", pch = 19, col = "green",  xlab = "Number of players (one mobs
points(N,PSI[2:50], type = "p", pch = 18, col = "black")
q <- expression(K, psi)
legend(40, 0.65, legend=q,
       col=c("green", "black"), pch= c(19,18), cex=0.8, box.lty=0)
```