I've considered about using time series models. But there are three problems that make this kind of models less practical.

- Handling missing data.

Since the missing data is the result of many stochastic behavior (server receiving update, users cleaning hard drive, the missing mechanism in the problem can't be handled by interpolation. If we just abandon the missing data, the sample size can be quite small to make accurate time series models.

- Identifying the model

Traditionally, we use acf/pacf plot first to roughly identify the order of ARIMA(p, r, q) model. Then, we will try some combinations of p, q, r to fit the model and use the order that produce lowest AIC to finally get the order. This is computationally consuming, which will make the speed of the program quite slow.

- Making predictions

Usually dynamic prediction is preferred in time series analysis. This means, if we have 500 data points and want to make predictions using time series models based on the given data, we have to use the model based on the 500 points to predict the 501^{st} time point. If we are predicting the 502^{nd} time point, we need re-fit the model with the same order as the previous model using the original 500 data points and the predicted 501^{st} time point, then use the newly fitted model to predict the 502^{nd} time point. Using this method can make the variance inflation slower than static prediction method, where we use the model based on the original 500 data points to predict all the latter data points. However, this procedure is quite slow using R functions.

- Programming difficulty

We will use some R packages, some of which are not in-built R packages, when we fit the model and make predictions. It's also quite hard to write our own functions.