Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/18/20230992

Involvement of the variance in the UCB algorithm regarding risk aversion and the regret bound

Yijia Zeng

University of California, San Diego, Math Department, 9500 Gilman Dr, La Jolla, 92122, California, United States

y2zeng@ucsd.edu

Abstract. The classical Upper Confidence Bound (UCB) algorithm implemented overestimates of the true mean of reward distributions based on the sample mean and the number of times such arms were chosen to decide the best arm. In this case, the variances, the essential components of any distributions, of reward distributions are dismissed. Moreover, in real-world applications, arms with relatively high means and small variances sometimes are preferable to arms with the highest mean and large variance. Such concerns are considered risk-aversion in the Multi-Armed Bandits algorithm. Additionally, since the combination of the variance and mean could help estimate the range of the distribution, proper utilization of the sample variance might help people to construct a tighter upper confidence bound to perform the UCB algorithm, leading to a smaller regret. In this paper, the author investigates and summarizes the current algorithm regarding risk aversion based on the estimations and implementations of variances in constructing the upper confidence bound in the UCB algorithm.

Keywords: variance, UCB algorithm, risk aversion, Multi-armed Bandit.

1. Introduction

Multi-armed Bandit algorithms are applied in real-life problems in multiple fields, including healthcare, finance, dynamic pricing, recommender system, etc. [1], to help agents make the best decisions among multiple options. In clinical trials, it is always difficult to determine the amount of dosage of a certain medicine to patients since an individual's need for such medicine is highly variable [1]. Consequences caused by incorrect dosage could be devastating. Therefore, the initial dosage should be determined based on both the average and variability of the performance of the dosage. This example represents a type of problem where people try to avoid adverse consequences for every single trail. Since people tend to minimize the risk in this condition, this type of problem is called a risk-aversion problem, which aims to find the option with a relatively high expected value and a low variance. The quantitative definition for "high" and "low" could vary among different problems and given algorithms.

There are several different measurements of risk, including Conditional value-at-risk (CVaR) [2,3], Mean-deviation risk [2], Mean-Variance Paradigm [4], etc. Xu, Haskell, and Ye proposed an algorithm called Risk-Averse Lower Confidence Bound (RA-LCB), which is a general structure of an algorithm for applying different kinds of risk measurement [2]. They defined a risk measure ρ that has some specific properties. By calculating the LCB index based on this function ρ, the arm with the expected trade-off between the mean and variance could be found [2]. Both CVaR, which considers the expected value of the reward in the worst case, and the Mean-deviation risk approach could be applied to this algorithm [2], obtaining an upper regret bound of $O(\sqrt{(\log n/n)})$. Except for this generalized algorithm that could implement diverse measurements, some researchers specifically investigated the CVaR-UCB [3], generating a regret bound of $O(\log n/n)$, which is a significant improvement [3].

Except for these kinds of risk measurement, this paper mainly focuses on the direct implementation of the variance. The author summarizes and compares some innovative risk-aversion UCB algorithms constructed based on the combination of the mean and variance as criteria, including ϕ-UCB [5], MV-LCB [6], GRA-UCB [6], and ARA-UCB [6]. Utilization of variance is not only effective in analyzing risks but is also used to develop a new kind of tighter upper bound, which is proven to outperform the classical UCB algorithm. In this paper, UCB1-Tuned [7] and UCB-V [8] are mainly discussed.

2. Classical UCB algorithm

Before showing the modifications of the UCB algorithm regarding risk aversion and the tighter upper bound, the classical UCB algorithm would be briefly introduced. The UCB algorithm is a classical algorithm in the Multi-armed bandit problem, which aims to explore the best arm and minimize the regret in such a process, where the regret is the potential total loss of not selecting the optimal arm. This algorithm is based on the strategy of “optimism in the face of uncertainty” [9]. Each arm could be assigned with a UCB index, which is an overestimate of the true mean. The algorithm would select the arm with the greatest UCB index for each round. The UCB index of the given arm would approach its true mean by selecting it more times in the long run. The formula for the UCB index is shown below [9]:

$$\mathrm{UCB}_i(t-1,\delta)=\begin{cases}\infty & \text{if } T_i(t-1)=0\\ \hat{\mu}_i(t-1)+\sqrt{\dfrac{2\log(\frac{1}{\delta})}{T_i(t-1)}} & \text{otherwise}\end{cases} \tag{1}$$

While $T_i(t-1)$ denotes the number of arm_i chosen before the time point t, and δ denotes the probability that the UCB index is smaller than the true mean of the given arm, such δ would be small. Due to the value of δ, $\sqrt{\frac{2\log(\frac{1}{\delta})}{T_i(t-1)}}$ is supposed to be a large value. According to the above formula, the only goal of this algorithm is to select the optimal arm the most while minimizing the regret based on the sample mean and the number of selections of each arm. Further revised versions of the UCB algorithm would follow a similar logic and be constructed based on this classical version.

3. UCB algorithm-based risk-aversion algorithm

The Mean-Variance paradigm is a fundamental risk modeling paradigm. Such an idea is originally proposed by Harry Markowitz [4]. He illustrated a geometric relationship of the trade-off between the mean and variance, indicating that the chosen option should be in the range of an efficient combination of the mean and variance [4]. Addressing the trade-off between the mean and variance, such a paradigm inspires many inventions of risk-aversion algorithms. The following sections discussed algorithms that address the similar idea introduced in the Mean-Variance paradigm by using the linear combination of the mean and variance to construct the modified UCB algorithm to, therefore, select the best arm.

3.1. ϕ-LCB algorithm

LCB is a descendant of the UCB algorithm. Instead of choosing the arm with the highest upper confidence bound, LCB defined the arm with the lowest lower confidence bound as the optimal arm in each selection. Intuitively, the risk-aversion algorithm aims to choose the arm with the least risk, so that the lower LCB is preferable.

The risk measure in ϕ-LCB is defined as $R(x)=f(E[X],\mathrm{Var}(X))$. According to this expression, the risk measure considered in ϕ-LCB is a function that takes the sample mean and sample variance as a parameter that maps to a number∈ R. In this paper, various types of such function fare considered, including continuous and discontinuous functions [5]. The continuous function addresses the case of a linear combination of the mean and variance, and the discontinuous function addresses the case of setting an acceptable threshold for the variance [5]. Given that algorithms based on the Mean-Variance paradigm are the main focus, only the continuous function would be discussed in the following sections. In the ϕ-LCB algorithm, the LCB index would be an underestimation of the value $f(\hat{\mu}_{i,t}, \hat{\sigma}^2_{i,t})$, where $\hat{\mu}_{i,t} = \frac{1}{t}\sum_{s=1}^{t} X_{i,s}$, the sample mean of arm_i, $\hat{\sigma}^2_{i,t} = \frac{1}{t}\sum_{s=1}^{t}(X_{i,s} - \hat{\mu}_{i,t})^2$, the sample variance of arm_i, and t denotes the number of selections of the given arm [5]. The value generated by the function ϕ would be subtracted from $f(\hat{\mu}_{i,t}, \hat{\sigma}^2_{i,t})$ to satisfy the requirement for underestimation. Such ϕ is defined as [5]:

$$\phi(0) = 0$$

ϕ is a strictly increasing function;

$|f(x_2) - f(x_1)| \le \phi(\|x_2 - x_1\|_1)$ for all $x_1, x_2 \in D$.

Based on such properties and inequality addressed by the Chernoff-Hoeffding bound, researchers generated the following inequality, where δ denotes the probability that the LCB index is not an underestimation [5]:

$$\left|f(\hat{\mu}_{i,t}, \hat{\sigma}^2_{i,t}) - f_i\right| \le \phi(6\sqrt{\frac{\ln\frac{1}{\delta}}{2t}}) \quad (2)$$

With the above information, it is clear that with the probability $1 - \delta$, $f(\hat{\mu}_{i,t}, \hat{\sigma}^2_{i,t}) + \phi(6\sqrt{\frac{\ln\frac{1}{\delta}}{2t}}) \le f_i$, such that $f(\hat{\mu}_{i,t}, \hat{\sigma}^2_{i,t}) + \phi(6\sqrt{\frac{\ln\frac{1}{\delta}}{2t}})$ is highly likely to be a lower bound of f_i. Therefore, the LCB index for the ϕ-UCB algorithm is defined as:

$$a_t = \mathrm{argmin}_{i=1\ldots K}\left[f(\hat{\mu}_{i,T_i(t-1)}, \hat{\sigma}^2_{i,T_i(t-1)}) - \phi(6\sqrt{\frac{\ln\frac{1}{\delta}}{2t}})\right] \quad (3)$$

At the step t, the algorithm would select the arm with the minimum LCB index to find out the arm with minimum risk measures.

Denote the risk measure of arm_α as R_α and the best arm with the least risk measure as R_{i^*}, the regret for the algorithm is defined as $R_n = \sum_{t=1}^{n} R_{a_t} - \sum_{t-1}^{n} R_{i^*}$, where R_{a_t} indicates that arm_α is chosen at the step t [5].

Suppose $\delta = \frac{1}{n^2}$, a typical selection of δ, then, for $\Delta_i = f_i - f_{i^*}$, the suboptimality gap, it is shown that

$$R_n \le \sum_{i:\Delta_i>0} \frac{36\Delta_i}{(\phi^{-1}(\Delta_i/2))^2} \ln n + \sum_{i:\Delta_i>0} \Delta_i \quad (4)$$

More specific analysis of regrets would be depended on the construction of the continuous function f.

Such a generalized structure of the LCB algorithm for the risk measurement based on the continuous function with inputs as the mean and variance has an upper bound of $O(\log n/n)$. The continuous function with inputs as the mean and variance could also address the case when the risk measurement is defined as a linear combination of the mean and variance. Therefore, the idea of the Mean-variance paradigm would be applied in this algorithm by making f as a linear combination of two inputs.

3.2. MV-LCB algorithm

Unlike ϕ-LCB which could generally implement all continuous functions of the mean and variance, Mean-Variance Multi-arm Bandit(MV-LCB) directly addresses the idea of the Mean-variance paradigm [10].

The risk measurement defined in MV-LCB is $MV_i = \sigma_i^2 - \rho\mu_i$, where ρ is the coefficient of the absolute risk tolerance ρ. Intuitively, since a low risk and a high expected value are desirable, the MV should be as small as possible, indicating a low variance or a high expected value.

Denote the true MV of the given arm_i as MV_i, and the empirical MV of such an arm at the step s as $\widehat{MV}_{i,s}$, the distance between the true MV and the estimated MV at the step s is represented as $|\widehat{MV}_{i,s} - MV_i|$. Based on the Chernoff-Hoeffding inequality [10]:

$$P(|\widehat{MV}_{i,s} - MV_i| \geq (5+\rho)\sqrt{\frac{\log\frac{1}{\delta}}{2s}}) \leq 6nK\delta \quad (5)$$

where $\exists i = 1, \ldots K, s = 1, \ldots n$

As the value of δ would be really small, the above equations indicate that $\widehat{MV}_{i,s} - (5+\rho)\sqrt{\frac{\log\frac{1}{\delta}}{2s}}$ would be highly likely to be a lower confidence bound for MV_i. Therefore, based on such inequality, the LCB index at the step s for arm_i is defined as [10]:

$$B_{i,s} = \widehat{MV}_{i,s} - (5+\rho)\sqrt{\frac{\log\frac{1}{\delta}}{2s}} \quad (6)$$

Based on such expression, the algorithm would select the arm with the lowest LCB index and update the $\widehat{MV}_{i,s}$ for the selected arm in each round.

For an empirical definition, $\widehat{MV}_{i,t} = \hat{\sigma}_{i,t}^2 - \rho\hat{\mu}_{i,t}$. Suppose the arm selected by a learning algorithm A in the step n has the MV represented by $\widehat{MV}_n(A)$, the regret in the step n is defined as $R_n(A) = \widehat{MV}_n(A) - \widehat{MV}_{i^*,n}$, where $\widehat{MV}_{i^*,n}$ denotes the empirical MV for optimal arm_{i^*} [10]. To generate the expression for the true regret, the unobserved performance of the optimal arm is expressed as $Y_{i,t}$:

$$Y_{i,t} = \begin{cases} X_{i^*,t} & \text{if } i = i^* \\ X_{i^*,t'} \text{ with } t' = T_{i^*,n} + \sum_{j<i, j\neq i^*} T_{j,n} + t & \text{otherwise} \end{cases} \quad (7)$$

Define the corresponding mean and variance of this variable as:

$$\tilde{\mu}_{i,T_{i,n}} = \frac{1}{T_{i,n}}\sum_{t=1}^{T_{i,n}} Y_{i,t} \tilde{\sigma}_{i,T_{i,n}}^2 = \frac{1}{T_{i,n}}\sum_{t=1}^{T_{i,n}}(Y_{i,t} - \tilde{\mu}_{i,T_{i,n}})^2 \quad (8)$$

Based on such expression of real observation and the previous definition of the regret for the algorithm R, researchers further generate an inequality:

$$R_n(A) \leq \sum_{i\neq i^*} T_{i,n}\hat{\Delta}_i + \frac{1}{n^2}\sum_{i=1}^{K}\sum_{j\neq i} T_{i,n}T_{j,n}\hat{\Gamma}_{i,j}^2 \quad (9)$$

$\hat{\Delta} = (\hat{\sigma}_{i,T_{i,n}}^2 - \tilde{\sigma}_{i,T_{i,n}}^2) - \rho(\hat{\mu}_{i,T_{i,n}} - \tilde{\mu}_{i,T_{i,n}})$ and $\hat{\Gamma}_{i,j}^2 = (\hat{\mu}_{i,T_{i,n}} - \hat{\mu}_{j,T_{i,n}})$ since this actual regret depends on real observations of running the optimal arms, which is hard to consider before the actual implementation. This paper would focus on the pseudo-regret, a weaker notion of the regret that only considered optimal action in expectations [10]. Therefore, instead of considering $MV_{i_*,n}$ a general expected performance of the optimal arm, denoted as MV_{i^*}, would be considered regardless of the current step number. Based on the previous derivation of the real regret, the pseudo regret over n steps is defined as:

$$\tilde{R}_n(A) = \frac{1}{n}\sum_{i\neq i^*} T_{i,n}\Delta_i + \frac{2}{n^2}\sum_{i=1}^{K}\sum_{j\neq i} T_{i,n}T_{j,n}\Gamma_{i,j}^2 \quad (10)$$

For $\Delta_i = MV_i - MV_{i^*}$ and $\Gamma_{i,j} = \mu_i - \mu_j$, to find an upper bound for such regret, the expected number of selections for the suboptimal arm is needed to be expressed. Following the same process of regret analysis in the UCB algorithm [9], a rough simplification of the regret is expressed as [10]:

$$E[\tilde{R}_n(A)] \leq O(\frac{K}{\Delta_{min}} + K^2\frac{\Gamma_{max}^2 \log^2 n}{\Delta_{min}^4 n}) \text{ for K as a constant} \quad (11)$$

Based on this expression, the theoretical upper bound for MV-LCB is approximately $O(\log^2 n)$ when there is only one unique optimal arm. However, in further studies [11], a tighter upper bound $O\log n$ is provided for this MV-LCB algorithm when there is only one unique optimal arm [10].

3.3. GRA-UCB and ARA-UCB algorithm

The GRA-UCB algorithm is designed for rewards distributions that follow independent Gaussian distribution [6]. Though the idea of the Mean-Variance paradigm is still followed and the expression $MV = \sigma^2 - \rho\mu$ is still utilized to investigate the optimal arm with the most efficient combination of the mean and variance, it was not designed by adding a general overestimation term on the whole MV like MV-LCB. However, it is considered that the special property of Gaussian distribution makes a proper estimation of σ^2 and adds overestimation terms on the mean only [6].

Based on Hoeffding inequality, $P(|\hat{\mu}_n - \mu| \geq \sqrt{\frac{\log t}{T_i(t)}}) \leq \frac{1}{t^2}$ is generated, indicating that $\hat{\mu}_n + \sqrt{\frac{\log t}{T_i(t)}}$ is highly possible to be an overestimate of μ as t increases. Therefore, by using the term $\hat{\mu}_n + \sqrt{\frac{\log t}{T_i(t)}}$ instead of μ, an underestimation of the MVwould be generated. Then, researchers estimate σ^2 based on the x^2 distribution.

Given that $P(\sigma^2 \leq \frac{(n-1)s_n^2}{x_{1-\alpha,n-1}^2}) = 1 - \alpha$, researchers use $\frac{(n-1)s_n^2}{x_{1-\alpha,n-1}^2}$ as an estimation for σ^2. Therefore, the UCB index for such an algorithm is defined as [6]:

$$B_i(t) = \frac{(T_i(t)-1)s_i^2(t)}{x_{a-\alpha,T_i(t)-1}^2} - \rho(\hat{\mu}_n + \sqrt{\frac{\log t}{T_i(t)}}) \quad (12)$$

Therefore, the algorithm is implemented by selecting the minimum index for each round. AGA-UCB is formed based on a similar logic by constructing asymptotic distribution of the empirical variance for a large n, such that [6]:

$$\sqrt{n}(s_n^2 - \sigma^2) \xrightarrow{d} N(0, \mu_4 - \sigma^4) \quad (13)$$

Without the assumption of reward distribution, this algorithm could be applied to any reward distribution. However, the justification for the selection of estimators for σ^2 is not clear for both algorithms. Given that estimators are overestimations of σ^2, the generated UCB index is not necessarily underestimation of the true MV. The pseudo-regrets for GRA-UCB and ARA-UCB are in similar forms to that of MV-LCB. Therefore, the regrets for both algorithms are bounded by $O(\log(N))$ [6].

3.4. A comparison of risk-aversion algorithms

According to the previous summaries for all these risk-aversion algorithms based on the estimation of the variance of risk distribution, it seems that each method utilized different ways to construct the LCB index. As a general algorithm that could consider all the continuous function involves μ and σ, ϕ-UCB [5] constructs the LCB index by defining a strictly increasing function and subtracting the outcome of such function to achieve the goal of underestimating the outcome of such a continuous function. Even though MV-LCB [10], GRA-UCB [6], and ARA-UCB [6] all address the idea of the MV paradigm [4] directly, MV-LCB constructs the LCB index by applying Hoeffding inequality on both the mean and variance, whereas GRA-UCB [6] and ARA-UCB [6] only apply such inequality on the overestimation of μ. For the estimation of σ^2, GRA-UCB [6] and ARA-UCB [6] consider using the reward distribution. However, the selection of the estimator needs further justifications.

For the regret analysis, all of the above algorithms have an upper bound of $O(\log n)$, indicating that they have comparable excellent performance in this task for a single optimal arm. Therefore, the selection of algorithms could depend more on the context instead of their efficiency.

DOI: 10.54254/2755-2721/18/20230992

4. Variance-based upper bound UCB algorithm

To further discover the usage of the variance of reward distribution in the UCB algorithm, this paper briefly investigates the algorithms UCB1-TUNED [7] and UCB-V [8], which shows that the upper confidence bound index involving the variance performs better than the classical UCB index.

4.1. UCB1-TUNED algorithm

UCB1-TUNED is a revised version of UCB1, where the structure of UCB1 follows that of the classical UCB algorithm [9]. The UCB index defined in UCB1 is $\hat{\mu}_i + \sqrt{\frac{2\log n}{T_i(t)}}$.

Instead of using the upper bound for UCB1, researchers defined an upper bound for the variance of each arm [7]:

$$V_i(T_i(t)) = \left(\frac{1}{T_i(t)}\sum_{s=1}^{T_i(t)} X_{i,s}^2\right) - \hat{X}_{i,T_i(t)}^2 + \sqrt{\frac{2\log t}{s}} \tag{14}$$

Given that the upper bound of the variance for a Bernoulli random variable is $\frac{1}{4}$, the UCB index for UCB1-TUNED is defined as $\hat{\mu}_i + \sqrt{\frac{\log n}{T_i(t)}\min(1/4, V_i(T_i(t)))}$. According to the experiment conducted in this paper, it can be concluded that the UCB1-TUNED performs substantially better than UCB1 [7]. However, there is no valid evidence that shows UCB1-TUNED has a theoretically better regret bound than UCB1($O(\log n)$) [7].

4.2. UCB-V algorithm

By applying the result of Freedman, a highly possible inequality can be obtained:

$$\left|\hat{X}_s - \mu\right| \leq \sqrt{\frac{2V_s x}{s}} + \frac{3bx}{s} \tag{15}$$

where V_s is the sample variance for given reward distribution, and x would be defined as ϵ, which is an exploration function that is non-decreasing as t increases.

Therefore the UCB index for this algorithm is defined as [8]:

$$B_{k,s,t} = \hat{X}_{k,s} + \sqrt{\frac{2V_{k,s}\epsilon_{s,t}}{s}} + c\frac{3b\epsilon_{s,t}}{s} \tag{16}$$

Where s denotes the number of selections of arm_k, and t denotes the number of total selections. Based on regret analysis, the regret of UCB-V is the upper bounded by $O(\log n)$ [8]. According to the computer experiments, for the problem with a high value of riskVaR, $Rn = \inf\{r: P(Rn \geq r) \leq \alpha\}$, the UCB index for a sub-optimal arm in UCB-V will converge significantly faster than that of UCB1 due to the consideration of variance [8]. Moreover, experiment results also show that due to such faster convergence, the UCB-V tends to result in a smaller regret. However, the opposite performance appears when the VaR is close to zero.

5. Conclusion

Based on the above analysis of the implementation of the variance in both constructing the risk-aversion UCB algorithm and the upper confidence bound index in the UCB algorithm, it seems that the variance of reward distribution has great potential in further investigation and development of the UCB algorithm.

For the risk-aversion algorithm, the notation from the algorithm is not that intuitive to apply to real life. Moreover, the coefficient of mean would affect the definition of an efficient combination of the mean and variance, which should be selected carefully. Since risk aversion is a new field in the UCB algorithm, investigations about the setting of coefficients in different kinds of problems could be done to further apply such an algorithm to more fields.

DOI: 10.54254/2755-2721/18/20230992

Since the variance represents the degree of deviation of data from the mean, the variance could be applied to generate an upper confidence bound index. With an implementation of the variance, it is possible to generate a tighter upper bound, which helps people converge to a true mean faster under any circumstances. Therefore, further research about using the variance to construct a better UCB index could be done.

Overall, this paper gathered and summarized the current implementation of the variance in the UCB algorithm, emphasizing the important role of the variance in the improvement of the UCB algorithm and shedding light on potential investigations of topics related to the variance and UCB index.

References

[1] Bouneffouf, D. and Rish, I. (2019). A survey on practical applications of multi-armed and contextual bandits. arXiv. DOI:10.48550/arXiv.1904.10040.

[2] Xu, J., Haskell, W. B. and Ye, Z. (2018). Index-Based Policy for Risk-Averse Multi-Armed Bandit. arXiv. DOI:10.48550/arXiv.1809.05385.

[3] Khajonchotpanya, N., Xue, Y. and Rujeerapaiboon, N. (2021). A revised approach for risk-averse multi-armed bandits under cvar criterion. Operations Research Letters 49(4), 465–472. https: //doi.org/10.1016/j.orl.2021.05.005.

[4] Markowitz, H. (1952). Portfolio selection. The Journal of Finance 7(1), 77–91. Accessed 2023-04-05.

[5] Zimin, A., Ibsen-Jensen, R. and Chatterjee, K. (2014). Generalized risk-aversion in stochastic multi-armed bandits. Computer ence. DOI:10.48550/arXiv.1405.0833.

[6] Liu, X., Derakhshani, M., Lambotharan, S. and Schaar, M. (2021). Risk-aware multi-armed bandits with refined upper confidence bounds. IEEE Signal Processing Letters 28, 269–273. https: //doi.org/10.1109/LSP.2020.3047725.

[7] Auer, P., Cesa-Bianchi, N. and Fischer, P. (2002). Finite-time analysis of the multi-armed bandit problem. Machine Learning 47, 235–256. https://doi.org/10.1023/A:1013689704352.

[8] Audibert, J.-Y., Munos, R. and Szepesv´ari, C. (2009). Exploration–exploitation tradeoff using variance estimates in multi-armed bandits. Theoretical Computer Science 410(19), 1876–1902. https://doi.org/10.1016/j.tcs.2009.01.016.

[9] Lattimore, T. and Szepesv´ari, C. (2020). Bandit Algorithms. Cambridge University Press, New York.

[10] Sani, A., Lazaric, A. and Munos, R. (2013). Risk-aversion in multi-armed bandits. Icml Workshop, 3275-3283. DOI:10.48550/arXiv.1301.1936.

[11] Vakili, S. and Zhao, Q. (2016). Risk-averse multi-armed bandit problems under mean-variance measure. IEEE Journal of Selected Topics in Signal Processing, 10(6), 1093-1111. DOI:10.1109/JSTSP.2016.2592622.