Classification of LHC Beam Loss Spikes using Support Vector Machines

G. Valentino[1,2], R. W. Assmann[1], R. Bruce[1] and N. Sammut[1,2]

[1] CERN, Geneva, Switzerland
[2] University of Malta, Msida, Malta
gianluca.valentino@cern.ch, ralph.assmann@cern.ch, roderik.bruce@cern.ch, nicholas.sammut@um.edu.mt

Abstract—**The CERN Large Hadron Collider's (LHC) collimation system is the most complex beam cleaning system ever designed. It requires frequent setups to determine the beam centres and beam sizes at the 86 collimator positions. A collimator jaw is aligned to the beam halo when a clear beam loss spike is detected on a Beam Loss Monitor (BLM) downstream of the collimator. This paper presents a technique for identifying such clear loss spikes with the aid of Support Vector Machines. The training data was gathered from setups held during the first three months of the 2011 LHC run, and the model was tested with data from a machine development period.**

Keywords- LHC Collimation System, Loss Spike Classification, Support Vector Machines
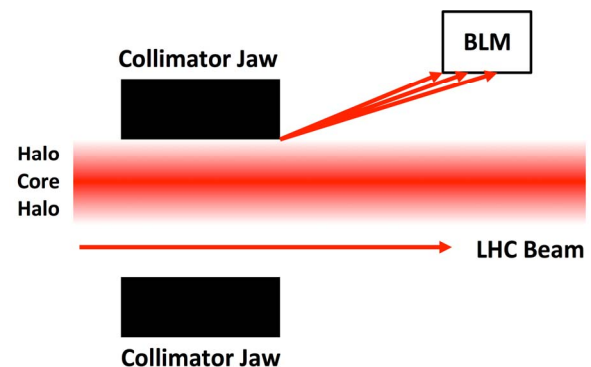

Figure 1. Scraping of the beam halo when the collimator jaw is aligned to the beam. The scattered secondary particles are detected by a BLM downstream and result in a time-varying signal.

I. INTRODUCTION

A. The LHC Collimation System

The CERN Large Hadron Collider (LHC) is a state-of-the-art particle accelerator that will collide nominal beams at 7 TeV [1], and is presently operating at 3.5 TeV. Its collimation system is designed to clean the beam halo and prevent any beam loss into the cold aperture, which could cause destructive quenches in the LHC's superconducting magnets [2]. Each collimator consists of two carbon or tungsten jaws, one on each side of the beam. Each jaw can be independently positioned with four degrees of freedom around the beam. The 86 ring collimators (43 per beam) are arranged in a three-fold hierarchy in order to capture primary, secondary and tertiary beam halos. According to their cleaning capabilities, they are located in a particular Insertion Region (IR) in the LHC tunnel at a certain distance in units of beam σ from the beam [3, 4].

B. Collimator Setup

During normal operation, it is important to maintain the correct hierarchy to guarantee the required level of machine protection. Due to unpredictable drifts in the beam orbit over a few months of operation, regular beam-based alignments or setups of the LHC collimators are necessary. The alignments allow the beam centres and beam sizes at each collimator to be determined, from which the hierarchy settings are calculated. Collimator setups can take place at 450 GeV (injection energy) or 3.5 TeV (at flat top, after squeeze or in collisions). Currently, the time needed for a full collimator setup is a limit on the frequency of setups, which in turn places constraints on the minimum β^* and luminosity reach of the LHC [5]. The luminosity is a measure of the number of particle collisions per unit time, and is an important parameter for a particle accelerator.

An automation of the collimator setup procedure can therefore allow for more frequent alignments, thus relaxing these constraints. A collimator jaw is defined to be aligned to the beam if a jaw movement towards the beam produces a clear beam loss signal. This loss signal is detected in an assigned Beam Loss Monitor (BLM) [6] located further downstream (see Fig. 1). A semi-automatic alignment algorithm that stops the jaw movement when the BLM signal exceeds a pre-defined threshold is discussed in [7].

C. Motivation for Spike Classification

Currently, a collimator expert is required to visually judge if a loss pattern is a clear indication that the jaw has touched the beam during the setup process. This is carried out when the jaws stop moving after the pre-defined beam loss threshold is exceeded. An example of an optimal (clear) loss spike is illustrated in Fig. 2, while a non-optimal loss spike is presented in Fig. 3. Previous work on beam loss signal classification was performed in the spatial domain in an attempt to classify patterns of losses all around the LHC ring [8]. However, for the collimator setup application the time-varying nature of the signal must be considered.
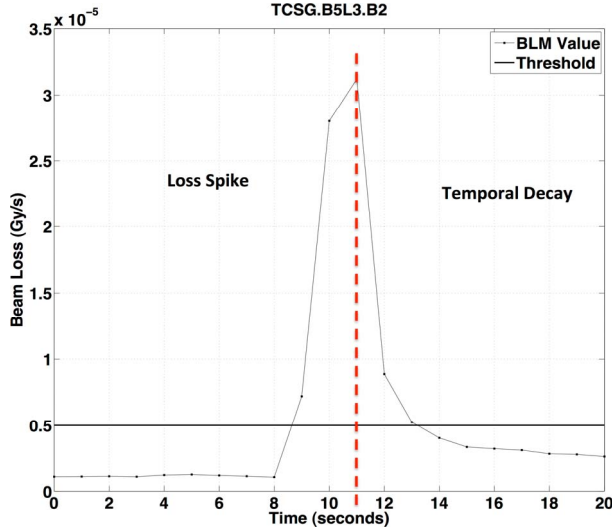
Figure 2. Optimal beam loss spike observed on the TCSG.B5L3.B2 collimator. The two components of the loss signal are clearly distinguishable.
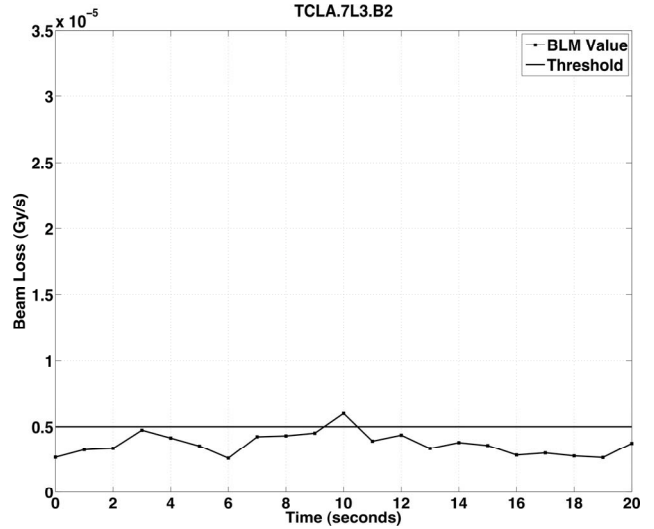

Figure 3. Non-optimal beam loss spike observed on the TCLA.7L3.B2 collimator. There are no clear-cut components although the losses have exceeded the threshold.

In fact, the beam loss signal (see Fig. 2) that is observed when a jaw touches the beam is the product of two physical processes. The first part of the signal is the loss spike. This sharp increase in the beam losses registered by the BLM is due to the scraping of particles from the beam halo, as illustrated in Fig. 1. The secondary particles formed as a result of the scraping are scattered into the BLM, and ionize the chamber to produce the spike. After the spike, the losses gradually decay to a "background" signal by means of a process that is still under study. Figure 2 illustrates an optimal loss spike, as the sharp increase and quick decay of the losses is a clear indication that the jaw has touched the halo.

II. SUPPORT VECTOR MACHINES

The Support Vector Machine (SVM) algorithm is a supervised learning technique that can be used for classification of data. It operates by maximizing the margin between the training data points and the decision boundary [9]. SVMs perform well and have been employed in many applications, ranging from image recognition to bioinformatics. Kernel functions are particularly useful when the data are non-separable. They map data on to a higher dimensional space so that a linear classifier can then be used. Examples include the linear, polynomial and Radial Basis Function (RBF) kernels. The RBF kernel is the most popular of these, as it is perceived to classify data with the best performance for most applications [10]. It maps data to an infinite-dimensional Hilbert space through the following relation:

$$K(x,y) = e^{\left(-\gamma \|x-y\|^2\right)} \qquad (1)$$

where x is the set of labelled training vectors, y is the set of classes and γ determines the width of the RBF. Another parameter is the penalty factor C, which if set too high results in a high penalty for non-separable data (overfitting), and if set too low leads to underfitting.

III. BEAM LOSS SPIKE CLASSIFICATION

A. Feature Selection

The performance of SVM is very dependent on the selection of the features. Fits can be applied separately to both components of the loss pattern in Fig. 2. The "loss spike" component was folded about the maximum value so that a Gaussian function could be fitted to it (see Fig. 4), while a power function was used to fit the "temporal decay" component as shown in Fig. 5. The fits were performed using the Ezyfit MATLAB tool [11], which uses MATLAB's built-in *fminsearch* function based on the Nelder-Mead method. A total of 6 input features were then considered, of which two pertain to the Gaussian fit and two to the power fit:

1. **Maximum Value:** determined by taking the maximum of the ten BLM values observed after the jaws have stopped moving. The losses may continue to increase after the threshold is exceeded, as shown in Fig. 2.

2. **Minimum Average:** the average of the three smallest loss points of the seven loss points immediately preceding the *maximum value*. The smallest values are considered to eliminate any spikes due to a previous movement. An optimal spike generally has a high *maximum value* relative to the *minimum average*.
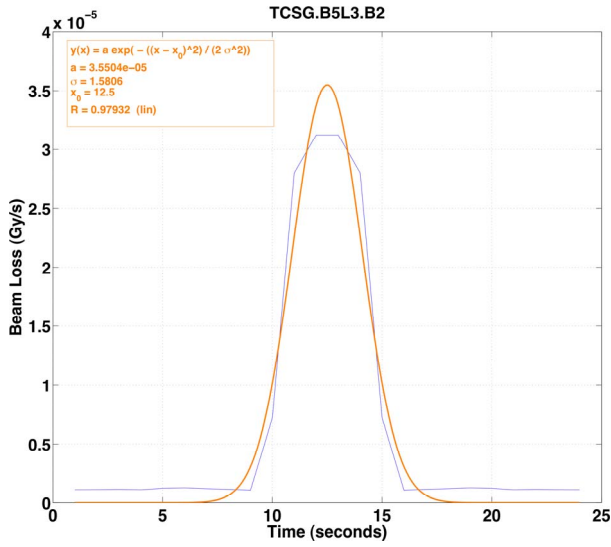
Figure 4. Fitting a Gaussian function to the loss spike folded about the maximum value.
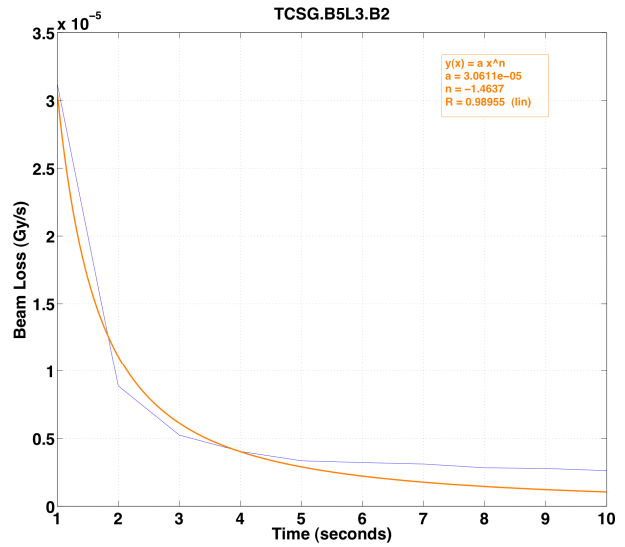

Figure 5. Fitting a power function to the temporal decay from the maximum value onwards.

3. **Variance:** the width of the Gaussian fit. Generally, a spike with a smaller width is more optimal, as it reflects a sharp increase and a quick decrease of the losses. This value is equivalent to σ for the example shown in Fig. 4.

4. **Gaussian Correlation Coefficient:** indicates the proximity of the loss pattern to the Gaussian fit. The closer this value is to unity, the sharper the loss spike. This value corresponds to R in Fig. 4.

5. **Power Coefficient:** a steep temporal decay is indicative of an optimal spike, equivalent to n in the example in Fig. 5.

6. **Power Correlation Coefficient:** indicates the proximity of the loss pattern to the power fit. The temporal decay becomes smoother as this value approaches unity. This value corresponds to R in Fig. 5.

Ten examples of the un-scaled data points, five from each class, are listed in Table I. Optimal spikes belong to the "+1" class, while non-optimal spikes belong to the "-1" class.

B. SVM Training

The LIBSVM tool presented in [12] was used for training and testing of the SVM model. The first step was to linearly scale the training data to values between -1 and 1. Scaling is done to avoid attributes in larger numeric ranges dominating those in smaller ranges. The scaling factors were then used to scale the test data in the same range.

The RBF kernel was chosen as it has less hyperparameters, and presents fewer numerical difficulties. A grid search on C and γ was performed using 5-fold cross-validation to determine the optimal values for these parameters. A value of 32768 was

TABLE I.
EXAMPLES OF UN-SCALED OBSERVATIONS AND THEIR FEATURE VALUES

Observation	Input Feature Values						Class
	1	2	3	4	5	6	
1	2.25118E-05	1.12315E-06	0.95979	0.98560	-1.09249	0.99182	+1
2	1.15883E-05	6.59159E-07	1.41285	0.98964	-1.28319	0.99160	+1
3	1.13180E-05	5.18305E-07	1.31249	0.99288	-1.06120	0.98716	+1
4	2.26029E-05	1.03017E-06	1.47528	0.97375	-0.91764	0.94713	+1
5	1.62120E-05	1.21613E-06	1.28399	0.96858	-1.01171	0.98796	+1
6	2.60442E-06	2.11741E-06	15.4612	0.09631	-0.28361	0.76970	-1
7	2.90270E-06	1.76021E-06	10.5298	0.45652	-0.08469	0.34159	-1
8	6.00977E-06	3.43113E-06	3.30181	0.89405	-0.28733	0.88192	-1
9	1.57701E-06	7.91728E-06	3.16919	0.93638	0.00318	0.01111	-1
10	1.16826E-06	5.98400E-07	3.80136	0.95359	-0.58202	0.95705	-1

Parameter	Values
Number of Features	6
Number of Classes	2
C parameter search range	$2^{-50} - 2^{50}$
γ parameter search range	$2^{-15} - 2^{15}$
C	32768
γ	0.125
ν	5
Training dataset prediction rate	97.2973 %
Testing dataset prediction rate	82.4324 %
Overall prediction rate	89.8649 %
Type of SVM	C-SVM
Kernel	RBF

obtained for C, while a value of 0.125 was determined for γ. Training was performed with these parameters using the RBF kernel to produce a model.

IV. EXPERIMENTAL RESULTS

A total of 444 samples were available, of which 222 were taken from a setup at 3.5 TeV for the tertiary collimators, and the remaining 222 samples were obtained from a setup at 3.5 TeV of the IR3 collimators. The numbers of samples for each category were chosen to maintain a 1:1 training to testing ratio. Predictions were made using the model developed in the training phase. The SVM model parameters, together with the prediction accuracy, are shown in Table II. An accuracy rate of 97.3% was achieved for the training data, while 82.4% of the test data points were classified correctly. This gives an overall prediction rate of 89.9%.

V. CONCLUSION

A novel method for distinguishing optimal beam loss spikes from non-optimal ones during the CERN LHC collimator setup was presented. Six input features specific to the classification problem were defined and fed into a support vector machine with a RBF kernel. A high overall prediction rate of 89.9% was achieved. Efforts will continue to suppress the number of mis-classifications. The SVM-trained recognition of beam loss spikes will then be integrated into the operational collimation software as part of a transition to a fast automatic collimator setup routine, which will significantly speed up an otherwise lengthy procedure.

REFERENCES

[1] L. Evans, "The Large Hadron Collider", New Journal of Physics vol. 9, p. 335, 2007.

[2] The LHC Design Report, Vol. 1 Chapter 18. Technical Report, CERN, 2004-003.

[3] R. W. Assmann, "Requirements for the LHC Collimation System", in Proceedings of the 8th European Particle Accelerator Conference, Paris, France, 2002, pp. 197-199.

[4] D. Wollmann et al., "First Cleaning with LHC Collimators", in Proceedings of IPAC'10, Kyoto, Japan, 2010, pp. 1237-1239.

[5] R. Bruce, R. W. Assmann, W. Herr, "Calculation Method for Safe Beta*", in Proceedings of IPAC'11, San Sebastian, Spain, 2011, pp. 1828-1830.

[6] E. B. Holzer et al., "Beam loss monitoring system for the LHC", in Proceedings of the IEEE Nuclear Science Symposium Conference Record, pp. 1052-1056.

[7] G. Valentino, R. W. Assmann, S. Redaelli, N. Sammut, D. Wollmann, "Semi-Automatic Beam-Based Alignment Algorithm for the LHC Collimation System", in Proceedings of IPAC'11, San Sebastian, Spain, 2011, pp. 3768-3770.

[8] A. Marsili, E. B. Holzer, P. Puzo, "LHC Beam Loss Pattern Recognition", in Proceedings of IPAC'11, San Sebastian, Spain, 2011, pp. 1353-1355.

[9] B. Boser, I. M. Guyon, V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers", in Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory, 1992, pp. 144-152.

[10] I. Steinwart, A. Christmann, "Support Vector Machines", New York: Springer-Verlag, 2008, p. 290.

[11] Frederic Moisy, software available at http://www.fast.u-psud.fr/ezyfit/

[12] C.-C. Chang and C.-J. Lin, ACM Transactions on Intelligent Systems and Technology 2, 27:1 (2011), software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.